Title: The estimation and compensation of processes with time delays

Candidate: Aidan O' Dwyer, B.E., M.Eng., C. Eng. M.I.E.I.

Ph. D. thesis

University Name: Dublin City University

Supervisor: Dr. John Ringwood

School: School of Electronic Engineering

Month and Year of Submission: August 1996

I hereby certify that this material, which I now submit for assessment on the program of study leading to the award of Ph.D. is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: Aidan O'Dwyer I.D. No.: 92701086

Date: 1 October 1996

Acknowledgements

I would like to express my gratitude to my supervisor, Dr. John Ringwood, for his encouragement and advice over the past four years. Thanks, John.

I would also like to thank my friends at the Dublin Institute of Technology, Kevin St., for their interest in my work; in addition, I would like to gratefully acknowledge the partial financial support of the European Social Fund Training for Trainers program, administered through the D.I.T.

Contents

Chapter 4: Frequency domain methods of model parameter and time delay estimation 106

Chapter 5: The compensation of processes with time delay 176

List of Figures

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Title: The estimation and compensation of processes with time delays

Author: Aidan O' Dwyer

ABSTRACT

The estimation and compensation of processes with time delays have been of interest to academics and practitioners for several decades. A full review of the literature for both model parameter and time delay estimation is presented. Gradient methods of parameter estimation, in open loop, in the time and frequency domains are subsequently considered in detail. Firstly, an algorithm is developed, using an appropriate gradient algorithm, for the estimation of all the parameters of an appropriate process model with time delay, in open loop, in the time domain. The convergence of the model parameters to the process parameters is considered analytically and in simulation. The estimation of the process parameters in the frequency domain is also addressed, with analytical procedures being defined to provide initial estimates of the model parameters, and a gradient algorithm being used to refine these estimates to attain the global minimum of the cost function that is optimised. The focus of the thesis is subsequently broadened with the consideration of compensation methods for processes with time delays. These methods are reviewed in a comprehensive manner, and the design of a modified Smith predictor, which facilitates a better regulator response than does the Smith predictor, is considered in detail. Gradient algorithms are subsequently developed for the estimation of process parameters (including time delay) in closed loop, in the Smith predictor and modified Smith predictor structures, in the time domain; the convergence of the model parameters to the process parameters is considered analytically and in simulation. The thesis concludes with an overview of the methods developed, and projections regarding future developments in the topics under consideration.

CHAPTER 1

Introduction

1.1 Background to research work

The aim of the research is to review identification and compensation methods for processes with time delays, and to develop appropriate identification and compensation strategies. A time delay may be defined as the time interval between the start of an event at one point in a system and its resulting action at another point in the system. Time delays are also known as transport lags, dead times or time lags; they arise in physical, chemical, biological and economic systems, as well as in the process of measurement and computation. Specific examples of systems with time delay are defined by Latour *et al.* (1967), Sandoz (1987), Papageorgiou and Messner (1989), Gendron *et al.* (1993), Menhaj and Hagan (1994) and Igarashi *et al.* (1994), amongst others. The estimation of time delays also arises in signal processing applications, where a time delay is also known as a time difference of arrival (TDOA) between two signals; such a measurement arises in underwater tracking applications, biomedicine, geophysics, astronomy, acoustics, seismology and telecommunications (Silva (1987), Salt *et al.* (1993), Shen *et al.* (1993), Sheridan (1993), Laguna *et al.* (1994) and Webster (1994)). Generally speaking, in the signal processing applications, it is the estimation of 'pure' time delay that is required, rather than the estimation of time delays in the presence of other process parameters.

Process parameter and time delay estimation techniques may be broadly divided into off-line techniques and on-line techniques, with on-line estimation requiring recursive estimation in a closed loop environment. On-line estimation is often called adaptive parameter estimation or sequential parameter estimation (Ljung (1987)) or identification by means of computers in on-line operation with the process (Isermann (1991), Isermann *et al.* (1992)). The choice of identification method for parameter estimation depends on the purpose of the identification, which determines the model needed and the accuracy required; a trade-off exists between required accuracy and computational effort when choosing the identification method. The choice of identification method also depends on whether the process may be interrupted from its normal operation, or if identification must take place during normal closed loop operation.

Apparent time delays may result when high order processes are approximated by means of lower order transfer functions (Seborg *et al.* (1989)). The time delay estimated in these applications may be a combination of a 'pure' time delay and contributions due to high order dynamic terms in the process transfer function. Bentley (1989) gives a pneumatic transmission line as an example of such a system and models it as a first order lag plus time delay (FOLPD) model. As mentioned, the purpose of the identification determines the model required. Newell and Lee (1989) state that there is much debate (in process control circles) over how complex a model may reasonably be identified from experimental data; they suggest that this depends on the data quality available (i.e. if the data is corrupted by noise) and the analysis technique used. The authors suggest that a cautious approach is to identify a FOLPD model from the experimental data and that an optimistic approach is to identify a second order system plus time delay (SOSPD) model from the data. Other authors that consider appropriate modelling methods for real processes include Latour *et al.* (1967), Pollard (1971), Edgar *et al.* (1981), Smith and Corripio (1985), Morari and Zafiriou (1989), Papageorgiou and Messner (1989), Seborg *et al.* (1989), Hang and Chin (1991), De Carvalho (1993), Gendron *et al.* (1993), Hang *et al.* (1994b), Kotob *et al.* (1994), Readle and Henry (1994), Schei (1994), Smirthwaite *et al.* (1994) and Yang (1994); these authors work is examined in more detail by O'Dwyer (1996a). A conclusion from this work is that even if the process has no physical time delay, it may be possible to model such a (possibly high order) process by a low order model plus time delay; it also appears reasonable that either a FOLPD model (for overdamped processes) or a SOSPD model (for overdamped or underdamped processes) should be estimated, as either of these approximate process models appears to be sufficiently accurate for many applications. However, if *a priori* information on the process is available (such as the process order), the estimation of the full order model plus time delay may be indicated; in any case, the work in the thesis will concentrate on the identification of processes that are adequately modelled by a linear model with time delay.

1.2 Thesis layout

The thesis will deal with the identification of process model parameters, and time delay, in both open loop and closed loop environments, together with the compensation of such systems. Chapter 2 classifies and outlines the approaches to model parameter and time delay estimation that have appeared in the literature; both off-line and on-line estimation techniques are treated. Chapter 3 develops one such estimation technique, involving the use of a gradient method, to estimate the parameters in open loop, in the time domain. Chapter 4 discusses frequency domain methods of parameter and time delay estimation; one such method uses a gradient approach to estimate the parameters. The focus is broadened in Chapter 5, in which compensation methods for processes with time delays are discussed in detail; one such compensation method, that involves the design of a modified Smith predictor, is treated in Chapter 6. Time domain methods that facilitate the estimation of the parameters in closed loop (in a Smith predictor structure) using appropriate gradient techniques are described in Chapter 7. In each of these chapters, conclusions as to the efficacy of the techniques discussed are reached, and further work is suggested. The conclusions of the work are outlined in Chapter 8, followed by a glossary of essential terms used in the thesis and a list of references. Thirteen technical reports, written by the author, are referenced throughout the thesis; these reports provide supplementary details on the topics discussed.

1.3 Thesis contributions

Original work contained in this thesis includes the following topics:

(1) The comprehensive review of methods defined in the literature for model parameter and time delay estimation (Chapter 2).

(2) The formulation of the cost function (which is a function of the error between the process and the model in the time domain) with respect to the time delay variation between the process and the model, when a variety of polynomials are used as approximations for the time delay variation. This work is detailed in Chapter 3.

(3) The development of seven theorems that are concerned with the minimisation of the cost function, when the parameters of a first order discrete stable process with time

delay are being estimated, in open loop, in the time domain. The theorems involve the calculation of the cost function and, in most cases, the proving that the cost function is unimodal with respect to the process time delay (as the model time delay varies), under defined operating conditions, using the principle of induction. These theorems are proved in Chapter 3.

(4) The development of algorithms for process frequency response measurement, in both open loop and closed loop, that use the ratio of the Fourier transforms of the output and input to the process and use a power spectral density approach. This work is detailed in Chapter 4.

(5) The development of analytical and iterative techniques to estimate the parameters of an arbitrary order model plus time delay from the process frequency response (Chapter 4). A model order estimation technique is also developed.

(6) The full review of methods defined in the literature for the compensation of processes with time delays (Chapter 5).

(7) The development of a modified Smith predictor, that improves the regulator response (when compared to the Smith predictor), while ensuring approximately the same servo response. This work is reported in Chapter 6.

(8) The development of five alternative gradient algorithms to the algorithm defined by Marshall (1979) and Bahill (1983), for the updating of the model parameters and time delay, in closed loop, in the time domain (Chapter 7); two of the algorithms are based on a Gauss-Newton gradient implementation and three of the algorithms are based on a Newton-Raphson gradient implementation. Three theorems are developed that show that the cost function is unimodal with respect to the process time delay as the model time delay varies, when the parameters of a number of process models are to be estimated.

CHAPTER 2

Approaches to model parameter and time delay estimation

2.1 Introduction

A time delay has been defined in Chapter 1 as the time interval between the start of an event at one point in a system and its resulting action at another point in the system. This chapter of the thesis will discuss time delay estimation methods (together with model parameter estimation methods, where appropriate) that have been proposed in the published literature; these methods may be broadly classified into time domain and frequency domain techniques.

Time domain estimation methods will be treated first. A number of off-line estimation techniques are outlined, for single input, single output (SISO) and multi-input, multi-output (MIMO) model structures, in open loop and in closed loop. A full discussion of multiple model estimation techniques will then be carried out; these methods typically involve estimating a number of models, each with a different value of the time delay, and subsequently determining the most appropriate model. However, these methods tend to be computationally intensive. A number of on-line estimation techniques will subsequently be treated, followed by a discussion of gradient methods of parameter and time delay estimation; the latter methods may be implemented in either open loop or closed loop, and in either an off-line or on-line manner. The estimation of time delays in the absence of other process parameters is also reviewed; such techniques are normally associated with signal processing applications.

Frequency domain estimation techniques may be classified in a similar manner to time domain estimation methods. The use of the frequency domain, as a means of estimating the parameters and time delay of a process model, has a certain intuitive appeal, since the time delay contributes to the phase term but not the gain term of the frequency response. A complete discussion of frequency domain estimation methods (which include higher order spectral algorithms) is provided.

Other possibilities for estimation are subsequently debated such as the use of neural networks, the use of process order identification methods and the implementation of the estimation strategies in the delta domain.

In each of the sections of the chapter, comparisons between the methods reviewed and conclusions as to the applicability of various classes of methods will be drawn, as appropriate. General conclusions from the literature review will be drawn and approaches to time delay estimation where original work may be usefully done will be outlined. The discussion in this chapter is further detailed by O'Dwyer (1996a).

2.2 Time domain methods for parameter and time delay estimation

2.2.1 Off-line estimation methods

The off-line estimation techniques may be broadly divided as follows:
(a) Methods using known process parameters
(b) Experimental open loop methods
(c) Experimental closed loop methods
(d) Other methods.

2.2.1.1 Methods using known process parameters

Estimation methods that use known process parameters are based on calculating an estimate of the parameters of a low order model plus time delay from the known parameters of a high order process. The methods are largely "rule of thumb" based methods that are unsuitable for the estimation of time delays of unknown processes. The methods, considered in detail by O'Dwyer (1992) previously, will not be reconsidered in this chapter.

2.2.1.2 Experimental open loop methods

These methods are based on estimating the parameters (including the time delay) from appropriate data gathered during tests while the process is in open loop. Typically, the input to the process is in step or pulse form. One of the first such methods was described by Ziegler and Nichols (1942), in which the time constant and time delay of a FOLPD process model are obtained by constructing a tangent to the step response at its point of inflection. The intersection of the tangent with the time axis at the step origin provides an estimate of the time delay; the time constant is estimated by calculating the intersection of the tangent with the value of the steady state output divided by the model gain. Other such tangent and point methods for estimating the parameters of a FOLPD model are described by Cheng and Hung (1985) and De Carvalho (1993), among others. The method may also be used to determine the parameters of a SOSPD model; Smith (1957), Perlmutter (1965), Meyer *et al.* (1967), Csaki and Kis (1969), Sundaresan *et al.* (1978) and Huang and Clements (1982) describe such approaches. The major disadvantage of all these methods is the difficulty of determining the point of inflection in practice.

Some methods that eliminate this disadvantage use two points on the process step response, to estimate the FOLPD model parameters, such as those described by Sunderesan and Krishnaswamy (1978) and Cheng and Hung (1985), or use two, three or more points on the process step response, to estimate the parameters of a SOSPD model, such as those described by Huang and Clements (1982), Huang and Huang (1993), Huang and Chou (1994) and Rangaiah and Krishnaswamy (1994a), (1996). An alternative experimental method involves calculating the parameters of an appropriate model from the area under the step response output curve (Nishikawa *et al.* (1984), Arzen (1987)).

Experimental open loop tests are a straightforward method of calculating the model parameters; however, the parameters of a FOLPD model approximation, determined by using actual step response data, may vary considerably depending on the operating conditions of the process, the size of the input step change and the direction of the change, with these variations being usually attributed to process nonlinearities (Seborg *et al.* (1989)). Harris and Mellichamp (1980) declare that a major drawback of an approach that involves the introduction of a step change is that the process must be sufficiently disturbed by the change to obtain reasonably accurate dynamic information; such a disturbance may well force the process outside the region of (approximately) linear behaviour. Arzen (1987) points out that methods to determine the dynamics of a process by examining its response to a deterministic signal such as a step or pulse input are conditioned on no drastic disturbances influencing the process. The time scale of the process must also be known in advance in order to determine when the transient response has been completed. Morari (1988) makes the important point that the method of judging model quality by comparing the process step response to the model step response is not necessarily the best means of optimising the model quality from the point of view of control system design; the author shows that three processes that have practically identical open loop responses may behave very differently under feedback.

2.2.1.3 Experimental closed loop methods

These methods are based on estimating the parameters (including the time delay) of a model from appropriate data gathered during the closed loop operation of the controlled process. The data is generally obtained from the closed loop step response; the methods typically involve the analytical calculation of the parameters of an appropriate process model from output measurements (such as the steady state value of the response and the first and second peak response values), of a unity feedback closed loop system under proportional control. Typically, the time delay is approximated in an appropriate manner (Yuwana and Seborg (1982), Jutan and Rodriguez (1984), Lee (1989), Jutan (1989), Bogere and Ozgen (1989)), though this is not absolutely necessary (Sung *et al.* (1994)). Chen (1989) and Lee *et al.* (1990) calculate the ultimate gain and frequency of a unity feedback closed loop system under proportional control, from the step response, and use these measurements to calculate the parameters of an appropriate open loop model. Hwang (1993), Hwang and Tseng (1994) and Hwang and Shiu (1994) use a combination of the methods based on step response measurements, and measurements of the ultimate gain and frequency, to determine the best process model; the latter two papers also outline similar identification strategies in closed loop when a PI or PID controller is used. Hwang (1995) brings together this work by outlining methods for the identification of a SOSPD process model in closed loop, by using the P, PI or PID controllers, and applying either a step, pulse or impulse test input signal in setpoint. In a more recent application, Kavdia and Chidambaram (1996) use the method of Yuwana and Seborg (1982) to calculate the parameters of a FOLPD model for an unstable process.

Refinements to the published algorithms are possible, as detailed by O'Dwyer (1996a); however, as mentioned in this report, the robustness of many of the estimation methods to noise on the process response is questionable. One method for which this comment does not apply is the characteristic areas method of Nishikawa *et al.* (1984), in which the area under the step response output curve is used to calculate the model parameters.

2.2.1.4 Other methods

Other off-line estimation methods do not naturally fall into any of the categories described earlier. Examples of methods that may be used to estimate the parameters and time delay of a linear SISO model include the following:

(a) Approximating the time delay by a Laguerre polynomial and using the standard (off-line) least squares estimation method to identify the parameters of the resulting model (De Souza *et al.* (1987), (1988), Salgado *et al.* (1988)) and

(b) Defining a state space model for the process and implementing a maximum likelihood estimate for the process parameters and the time delay based on this parameterisation (Nagy and Ljung (1991)).

Representative methods that have been used to determine the parameters and time delay of a linear MIMO model are

(a) The extension of a method to estimate the model order, defined for SISO systems, which is based on inspecting the near singularity of the information matrix, to also estimate time delays, if the ranges of the time delays are known (Mancher and Hensel (1985)) and

(b) The resolving of output signals of MIMO processes into a set of independent output signals for SISO processes by using persistently exciting Walsh function input signals; the Walsh functions may then be used to estimate the parameters and the time delay of each of the SISO processes (Bohn (1985)).

2.2.2 Methods based on multiple model estimation

These methods are based on the estimation of a number of different process models, for different values of the time delay. The model parameters chosen are those that minimise a cost function that depends on the difference between the process and the model outputs. One of the best examples of the approach is given by Baur and Isermann (1978), who use recursive correlation analysis with least squares parameter estimation to detect $m_{max}(d_{max} - d_{min})$ separate models, where m_{max} = the maximum model order and the time delay index (which is the integer value of the time delay divided by the sample time) lies between d_{min} and d_{max}. A loss function V(m,d), based on the residuals, is minimised as model order is varied; the optimum estimates of model order, m, and time delay index, d, are determined if V(m+1,d) and V(m,d+1) do not decrease significantly in relation to V(m,d). Other authors that also estimate the model order, parameters and time delay index using a multiple model method include Gabay and Merhav (1976), Bokor and Keviczky (1984), Peterka (1989), Hemerly (1991), Musto and Lauderbaugh (1991), Warwick and Kang (1993) and Tuch *et al.* (1994). Some authors concentrate on estimating the time delay and process parameters only; the time delay is estimated by minimising the loss function as the time delay index is varied, with the process parameters estimated using other methods. Among the authors that discuss such techniques are Hsia (1969), Rao and Sivakumar (1979), Rao and Palaniswamy (1983), Hansen (1983), Pearson and Wuu (1984), Wuu and Pearson (1984), Cheng and Hung (1985), Abrishamakar and Bekey (1985), (1986), Batur (1986), Agarwal and Canudas (1987), Jiang (1987), Juricic (1987), Kim *et al.* (1987), Unbehauen and Rao (1987), Peter and Isermann (1988), Casted (1989), Co and Ydstie (1990), Zheng and Feng (1990), Ferretti *et al.* (1991), Schei (1992), Lublinsky and Fradkov (1993), Chen and Loparo (1993), Leva *et al.* (1994), Readle and Henry (1994), Ferretti *et al.* (1995) and Wang and Clements (1995).

The multiple model estimation technique may also be used to estimate the parameters of multiple-input, single output (MISO) or MIMO process models with time delays. Authors that estimate the model order, parameters and time delay indices using such methods for these applications include Blessing *et al.* (1978), Bokor and Keviczky (1984), Mancher and Hensel (1985), Xu (1989) and Haest *et al.* (1990).

The attraction of multiple model estimation methods is that the grid searching used will facilitate the estimation of the parameters corresponding to a global minimum of a cost function, even in the presence of local minima, provided enough models are estimated. The method is relatively crude compared to the use of gradient search methods (discussed in Section 2.2.4), and it is also more computationally intensive; however, the latter methods do not guarantee the estimation of the parameters corresponding to the global minimum, in the presence of local minima.

2.2.3 On-line estimation methods

On-line time delay estimation requires recursive estimation of the time delay in a closed loop environment. The techniques may be classified as follows:

(a) Methods that use rational approximations for the time delay, followed by recursive identification of the model parameters and

(b) Methods that involve overparameterisation of the process in the discrete time domain.

2.2.3.1 Methods that use rational approximations for the time delay

The following rational approximations may be used for the time delay:

(a) The Taylor's series expansion
(b) The Pade approximation
(c) The Laguerre approximation
(d) The Product approximation (or Paynter delay line)
(e) The direct frequency response approximation technique
(f) The Bessel approximation
(g) A transfer function approximation (from Marshall (1979)) and
(h) Numerical optimisation (e.g. the equiripple formula); this is defined by Piche (1990).

These approximations have been detailed by O'Dwyer (1996a); for example, the first order Taylor's series approximation for the time delay, $e^{-s\tau}$, is $1 - s\tau$.

Seborg *et al.* (1989) declare that when the time delay is less than one tenth of the time constant (in a FOLPD process model structure), then a first order Pade approximation for the time delay is accurate to within engineering accuracy, considering that most processes behave like low pass filters; correspondingly, the second order Pade approximation is accurate to within engineering accuracy when the time delay is less than one fifth of the (repeated) time constant of a more general process model structure.

When the time delay is approximated by a rational polynomial, the resulting model parameters are normally estimated in a discrete time environment using an algorithm based, for instance, on recursive least squares (RLS); the time delay may then be deduced from the model parameters identified. Such an approach is outlined by Roy *et al.* (1990), (1991a), (1991b), (1991c), (1993a), (1993b), Boje and Eitelberg (1991), Bai and Chyung (1993), Fernandes and Ferriers (1994) and Yasterbov and Grzywaczewski (1994). However, the method defined by Roy *et al.* (1990), (1991a), (1991b), (1991c), in which the time delay is modelled by a zero in the continuous time domain, with the parameters of the model being identified using the RLS algorithm in the discrete time domain, did not work for simulations taken by Kelly (1991) or O'Dwyer (1992) i.e. it was not possible to estimate the time delay from the resulting process parameters identified.

2.2.3.2 Overparameterisation of the process in the discrete time domain

The method of overparameterisation involves subsuming the time delay term into an extended (or overparameterised) z domain numerator polynomial. The corresponding parameters are estimated using a recursive estimation scheme, and the time delay is calculated based on the parameters identified; for a noise free system, all numerator parameters whose indices are smaller than the time delay index should be identified as zero. Only values of the time delay that are integer multiples of the sample period are directly estimated by the method. The part of the time delay that is a fraction of the sample period may be calculated from the numerator parameters identified, for processes that can be modelled by a FOLPD model (O'Dwyer (1992), (1993)) and for processes that may be modelled by a SOSPD model (Thomson *et al.* (1989)); however, the robustness of these methods of estimation in the presence of noise is questionable.

Many overparameterisation methods have been defined to calculate the numerator (and denominator) parameters, and subsequently the time delay, for processes that may be modelled in SISO form or MIMO form. Kurz (1979) and Kurz and Goedecke (1981), for example, define a robust method for estimating the SISO model parameters that is equivalent to determining the best match between the impulse response of the overparameterised model and the impulse response of a non-overparameterised model with a pure time delay; the method suffers from the disadvantage of having a heavy computational load. Other methods offer various trade-offs between robustness and computational load, such as those described by Biswas and Singh (1978), Astrom and Zhou (1981), Friedlander (1982), Wong and Bayoumi (1982), Habermayer and Keviczky (1985), Habermayer (1986), Batur (1986), De Keyser (1986), Koivo *et al.* (1988), Hu *et al.* (1988), Keviczky and Banyasz (1988), Najim *et al.* (1988), Xu (1988), Teng and Sirisena (1988), Landau (1990), Teng (1990), Guez and Pioviso (1991), Lundh and Astrom (1994) and Readle and Henry (1994). Other authors describe a recursive method to estimate the parameters, order and time delay index for both a stochastic system and a deterministic system, using an overparameterised method to estimate the time delay (Chen and Zhang (1990) and Zhang and Chen (1990)). In an interesting paper, Keviczky and Banyasz (1992) identify the time delay index using overparameterisation in the delta domain (see Section 2.4).

Other authors identify MIMO process models (with time delays) using the method of overparameterisation; Gurubasavaraj and Brogan (1983), for instance, extend the method of Kurz and Goedecke (1981) to estimate the time delay for each input-output pair of a MIMO process. Simulation results presented by the authors show that the time delays may be estimated in 20 sample periods, for a 2x2 MIMO process with a maximum time delay index of 4; the process order is however assumed known *a priori*. Other authors that use overparameterisation for this application include Song and Xu (1985) and Zhang and Chen (1990).

The attractiveness of the method of overparameterisation as a means of estimating model parameters and time delay is that it is a natural extension of methods used in delay-free identification applications. However, the method has many disadvantages.

(1) The computational burden of the RLS algorithm increases with the square of the number of estimated parameters (De Keyser (1986), Glentis and Kalouptsidis (1992), Ferretti *et al.* (1995)).

(2) The persistent excitation condition (a condition for parameter convergence) is more difficult to satisfy for overparameterised models (Kim *et al.* (1987), Dumont *et al.* (1993)).

(3) The adaptive capability of the corresponding controller is degraded, as it takes a long time for the parameters to be retuned if a change in the process dynamics occurs (Kim *et al.* (1987)). However, it is possible by introducing a perturbation signal into the regressor vector, when the parameters of the model with delay are being estimated, to achieve a similar convergence rate for the parameters of an overparameterised model as for the parameters of a non-overparameterised model (Xia *et al.* (1987), Xia and Moore (1989)).

(4) The presence of a high order numerator polynomial increases the likelihood of common factors in the numerator and denominator polynomials in the estimation model, rendering identification more difficult (Dumont *et al.* (1993)).

(5) The overparameterisation method is not robust if a load disturbance is present, or if

measurement noise is significant (Lee and Hang (1985)). However, Xia *et al.* (1987) and Xia and Moore (1989) state that injecting an excitation signal into the regressor vector (for RLS estimation (1987) or recursive extended least squares (RELS) estimation (1989)) allows the parameters of a model of one order of overparameterisation (1987) or arbitrary degree of overparameterisation (1989) to have the same guaranteed convergence as the parameters of a non-overparameterised model (i.e. ill-conditioning is avoided for the overparameterised model) for both models with white noise excitation (1987) and coloured noise excitation (1989).

From this discussion, the biggest disadvantage of the overparameterisation method for the identification of a process with time varying delay in closed loop, perturbed by a pseudo-random binary signal (PRBS), is the extra computational burden associated with identifying a greater number of numerator parameters. In an attempt to reduce the computational burden associated with the overparameterisation method, the following ideas may be worth considering:

(a) If the time constants of the process do not change significantly, then the denominator parameters need not be estimated on-line; as well as reducing the computational burden in the estimation stage, other advantages of this scheme are that excessive fluctuation of the denominator parameters is avoided and the denominator parameter estimates cannot drift into or near an undesirable region. This suggestion was made by Vogel and Edgar (1982). A further suggestion made by Seborg *et al.* (1986) is that selective updating of certain model parameters be employed when the number of parameters of the process to be estimated is large; such selective updating could be achieved by only updating those parameters that give a significant improvement in the residual of the model fit.

(b) The sampling interval could be adapted to reduce the number of parameters to be estimated. To this end, Seborg *et al.* (1986) suggest that the sampling period be chosen so that the time delay index has a value of two or three; such slow sampling, the authors suggest, has the additional advantage of increasing the robustness of the corresponding adaptive controller. This advice may be relevant only for small values of the time delay as otherwise it may conflict with the most often quoted rule of thumb that the sampling period should be between one fifth and one fifteenth of the 95% rise time of the process step response (Isermann (1989)).

The methods of time delay estimation using the overparameterised model that appear most robust are those of Kurz (1979) and Teng and Sirisena (1988). For a practical application, the method of Teng and Sirisena (1988) seems to be most promising, because of its relative computational simplicity. The method of overparameterisation as a means of estimating time delays may be extended in application from SISO processes to MISO processes and MIMO processes. Surprisingly, the methods that have been well documented for the estimation of the process time delay in the SISO environment have not been widely applied to the identification of time delays in MIMO processes in the available literature (one of the few exceptions is the method outlined by Gurubasavaraj and Brogan (1983)). This topic is discussed more fully by O'Dwyer (1996l).

2.2.4 Gradient methods of parameter and time delay estimation

2.2.4.1 Introduction

Gradient methods of parameter estimation are based on updating the parameter vector (which includes the time delay) by a vector that depends on information about the cost function to be minimised. The gradient algorithms considered normally involve expanding the cost function as a second order Taylor's expansion around the estimated parameter vector. The cost function is given by

$$J(n) = 0.5\sum_{j=0}^{N-1} e^2(n-j) \tag{2.1}$$

with J(n) = cost function and e = error = process output minus model output. A second order Taylor's series expansion of the cost function may be determined from equation (2.1) to be

$$J(n+1) = J(n) + \frac{\partial J(n+1)}{\partial \theta(n)}(\theta(n) - \theta^*(n)) + 0.5(\theta(n) - \theta^*(n))\frac{\partial^2 J(n+1)}{\partial \theta^2(n)}\left(\theta(n) - \theta^*(n)\right) \tag{2.2}$$

with $\theta(n) \in \Re^n$, $\theta(n)$ = parameter vector and $\theta^*(n)$ = optimum parameter vector. An estimate of the parameter vector is determined by minimising $J(n+1)$ with respect to the parameter vector. A simplified updating strategy based on this minimisation is

$$\theta(n+1) = \theta(n) + \mu\phi(n) \tag{2.3}$$

with
$$\phi(n) = -\left[\frac{\partial^2 J(n)}{\partial\theta^2(n)}\right]^{-1}\left[\frac{\partial J(n)}{\partial\theta(n)}\right] \tag{2.4}$$

and with $\phi(n) \in \Re^n$ and μ = learning rate; the default value of μ = 1.0. The partial derivative of the cost function with respect to the parameter vector may be determined recursively (from equation (2.1)) to be

$$\frac{\partial J(n+1)}{\partial\theta(n)} = \frac{\partial J(n)}{\partial\theta(n)} + e(n+1)\frac{\partial e(n+1)}{\partial\theta(n)} \tag{2.5}$$

with the starting value of the partial derivative of the cost function with respect to the parameter vector assumed zero. The calculation of the second partial derivative of the cost function with respect to the parameter vector determines the nature of the optimisation algorithm. Ljung (1987) divides these optimisation algorithms into three classes:

(1) The updating vector is a function of the cost function, the partial derivative of the cost function with respect to the parameter vector and the second partial derivative of the cost function with respect to the parameter vector. The Newton-Raphson algorithm is an example; under these circumstances, the second partial derivative of the cost function with respect to the parameter vector (labelled the Hessian matrix), calculated using equation (2.5), is given by

$$\frac{\partial^2 J(n+1)}{\partial\theta^2(n)} = \frac{\partial^2 J(n)}{\partial\theta^2(n)} + e(n+1)\frac{\partial^2 e(n+1)}{\partial\theta^2(n)} + \frac{\partial e(n+1)}{\partial\theta(n)}\left[\frac{\partial e(n+1)}{\partial\theta(n)}\right]^T \tag{2.6}$$

(2) The updating vector is a function of the cost function and the partial derivative of the cost function with respect to the parameter vector; in this case, an estimate of the second partial derivative of the cost function with respect to the parameter vector is used. The Gauss-Newton algorithm (also called the method of scoring, the modified

Newton-Raphson algorithm or the quasilinearisation algorithm), the Levenberg-Marquardt algorithm and the steepest descent algorithm are examples; the second partial derivative of the cost function with respect to the parameter vector for the Levenberg-Marquardt algorithm is

$$\frac{\partial^2 J(n+1)}{\partial \theta^2(n)} = \frac{\partial^2 J(n)}{\partial \theta^2(n)} + \frac{\partial e(n+1)}{\partial \theta(n)} \left[\frac{\partial e(n+1)}{\partial \theta(n)} \right]^T + \delta I \tag{2.7}$$

with δ being a positive constant and the identity matrix, $I \in \Re^{nxn}$. The updating vector, $\phi(n)$, in this case is given by (from equations (2.4) and (2.7))

$$\phi(n) = \frac{-\left[\frac{\partial^2 J(n)}{\partial \theta^2(n)} \right]^{-1} \frac{\partial e(n+1)}{\partial \theta(n)} e(n+1)}{\lambda(n) + \left[\frac{\partial e(n+1)}{\partial \theta(n)} \right]^T \left[\frac{\partial^2 J(n)}{\partial \theta^2(n)} \right]^{-1} \left[\frac{\partial e(n+1)}{\partial \theta(n)} \right]} \tag{2.8}$$

and

$$\left[\frac{\partial^2 J(n+1)}{\partial \theta^2(n)} \right]^{-1} = \frac{1}{\lambda(n)} \left[\left[\frac{\partial^2 J(n)}{\partial \theta^2(n)} \right]^{-1} + \mu \frac{\phi(n)}{e(n+1)} \left[\frac{\partial e(n+1)}{\partial \theta(n)} \right]^T \left[\frac{\partial^2 J(n)}{\partial \theta^2(n)} \right]^{-1} \right] + \delta I \tag{2.9}$$

with $\lambda(n)$ = forgetting factor and $\theta(0)$ = known starting values.

The Gauss-Newton algorithm omits the addition of the δI term. These two algorithms are special cases of the Newton-Raphson algorithm in which the following conditions are fulfilled: (a) the partial derivative of the cost function with respect to the parameter vector is assumed to be zero at the current parameter vector (this is obviously an approximation, as the partial derivative will only be zero at the optimum parameter vector) and (b) the error multiplied by the second partial derivative of the error with respect to the parameter vector may be neglected (Soderstrom and Stoica (1989) state that this is valid close to the optimum parameter vector).

In all these cases, the starting value of the second partial derivative of the cost function with respect to the parameter vector is given as a multiple of the identity matrix.

The Hessian matrix for the steepest descent algorithm equals the identity matrix; the updating vector, $\phi(n)$, is given by equation (2.8), with the appropriate substitution.

(3) The updating vector is a function of the cost function only; these algorithms either form gradient estimates by difference approximations and proceed as in (for example) the Gauss-Newton algorithm, or have other specific search patterns.

Other gradient algorithms would not naturally fall into these classes; one example would be the least mean squares (LMS) algorithm defined by Widrow and Stearns (1985):

$$\theta(n+1) = \theta(n) - 2\mu(\partial e(n)/\partial\theta(n)) \tag{2.10}$$

The choice of the gradient algorithm for a particular application depends on the desired speed of tracking and the computational resources available. Draper and Smith (1981) declare that the Gauss-Newton algorithm combines the best features of the Newton-Raphson method and the steepest descent method, though the convergence of the algorithm is slower than that of the Newton-Raphson algorithm. The authors declare that the steepest descent method, though straightforward, often converges very slowly to the optimum parameter vector after rapid initial progress. Smith and Friedlander (1985) agree, declaring that while the recursive Gauss-Newton algorithm is quadratically convergent near a local minimum of the cost function, the steepest descent algorithm is only linearly convergent in the same situation. The Gauss-Newton algorithm has the advantage over the Newton-Raphson algorithm of being less computationally intensive; Ljung (1987) also states that the approximation used to determine the Gauss-Newton algorithm ensures that the Hessian matrix is positive semi-definite, which means that convergence is guaranteed to a stationary point. On the other hand, Soderstrom and Stoica (1989) declare that the convergence of the Newton-Raphson algorithm is quadratic, whereas in practice the convergence of the Gauss-Newton algorithm is linear but fast.

In a more general point about the use of gradient methods, it is important that the error surface in the direction of the time delay (and indeed the other parameters) should be unimodal. The existence of a multimodal error surface in the direction of the time delay has serious consequences for the use of the gradient algorithm; indeed, it appears that the task of determining a global minimum in the presence of local minima is a very knotty problem. Vanderplatts (1984) states that the estimation process must be started from various initial estimates to see if a consistent optimum may be obtained under these conditions; reasonable assurance is then felt that this optimum point corresponds to the true global minimum. Rekliatis *et al.* (1983) state that the only practical strategy for locating global minima in these situations is a method called "multistart with random sampling". This strategy involves multiple optimisation runs, each initiated at a different starting point. The starting points are selected by sampling from a uniform distribution. The global minimum is then the local minimum with the lowest cost function value among all the local minima that may be identified. Scales (1985) suggests that in practice, one usually has to assume that it is possible to make a guess at the position of the global minimum that is sufficiently good so that no extraneous local minima interfere with the minimisation process. Ferretti *et al.* (1996) declare that the use of a filter on the data increases the range of time delay over which the cost function is unimodal; the bandwidth of the filter is related to an initial estimate of the time delay uncertainty. However, the speed of convergence of any gradient algorithm used is reduced by the inclusion of a filter in this manner.

It may be possible to improve the chances that the global minimum of the error surface may be determined, even if the error surface is multimodal, by adapting techniques defined by Demuth and Beale (1977), amongst others, that improve backpropagation in neural networks. One technique defined by these authors is that of learning with momentum; the authors declare that momentum acts like a low pass filter on the error surface, allowing the possibility of sliding through local minima. The idea is that a change in the parameter will be put equal to the momentum constant (typically 0.95) times the previous change in the parameter plus 0.05 times the present change in the parameter. A further possibility defined is to use an adaptive learning rate; the authors propose that the learning rate should be decreased by a factor of 0.7 if the new error exceeds the old error by a factor of 1.04 and increased by a factor of 1.05 if the new error is less than the old error. However, no theoretical basis is given for these two suggestions. A number of other authors have defined adaptive learning algorithms; among them are Ho and Hsu (1992), Knapp and Wang (1992) and Qiu *et al.* (1992). These algorithms have been developed from a trial and error approach. Silva and Almeida (1990) and Sato (1991) also discuss the use of momentum and learning rate terms in the application.

On a practical level, since all of the gradient implementations may identify parameters corresponding to a local minimum rather than a global minimum, it is important to commence iterations at good initial values of the parameters, which may be obtained by physical insight for a physically parameterised model structure. A further advantage in starting off at good initial values is that the number of iterations required for good identification is lower and the total computing time required is less.

2.2.4.2 Gradient algorithms for estimation based on the Newton-Raphson, Gauss-Newton and steepest descent methods

A number of authors have defined gradient algorithms based on the Newton-Raphson method, for estimating process parameters; Liu (1990), for example, defines a parameter updating scheme for an n^{th} order process model plus time delay based on the algorithm. Other algorithms for appropriate parameter updating based on the Newton-Raphson approach include the method defined by Zhao and Sagara (1990).

The use of the Gauss-Newton algorithm to estimate process parameters was perhaps first proposed by Marshall (1979), who uses such an algorithm to identify the parameters of a FOLPD model, in a Smith predictor structure. A number of assumptions are made in this analysis; Bahill (1983) subsequently used these assumptions to facilitate the development of an equation for the required change in the model time delay as a result of the change in the process time delay. A number of modifications of the algorithms defined above have also been considered, including those implemented by Kaya and Scheib (1984) (who implement Marshall's (1979) scheme to update the time delay estimate, and estimate the parameters of a first order lag (FOL) model of the non-time delayed process using the RLS algorithm), Romagnoli *et al.* (1988) and O'Connor (1989). The Gauss-Newton algorithm has also been used in open loop to estimate process parameters; Durbin (1984a), (1984b), (1985), for instance, uses the algorithm to estimate the parameters of a FOLPD model of the process. Simulation results quoted by Durbin (1985) show that a change in time delay index from 10 to 14 is followed in about 25 samples, with a change in time delay index from 14 to 10 followed in about 40 samples. Other such gradient algorithms are defined by Wong (1980), Brewer (1988) and Banyasz and Keviczky (1988), (1994). Other authors, such as Smith and Friedlander (1985) and Pak and Li (1992), concentrate on estimating the time delay only using the algorithm.

The straightforward nature of the steepest descent algorithm has motivated a number of authors (such as Elnagger *et al.* (1989), (1990a), (1990b), (1991), (1992), (1993)) to apply it to the estimation of process parameters. These authors (1990a) estimate the non-delay parameters using the RLS algorithm, and estimate the delay using the steepest descent algorithm. Other authors, such as Robinson and Soudack (1970), concentrate on estimating the time delay only using the algorithm.

2.2.4.3 Other gradient algorithms for time delay estimation

As mentioned in Section 2.2.4.1, there are other gradient algorithms that may be used for model parameter and time delay estimation; Gawthrop and Nihtila (1985), for instance, estimate a pure time delay in a noise free environment by updating the time delay based on the partial derivative of the error squared with respect to the time delay. Gawthrop *et al.* (1989) use the same technique to estimate the parameters of a continuous time SISO process with time delay. Other algorithms of the type under discussion for estimating the model parameters and time delay are defined by Pupeikis (1985), Shah *et al.* (1988) (who use the LMS algorithm), Boudreau and Kabal (1992), (1993), Hwang and Chuang (1994) and Lim and Macleod (1995). Algorithms of this type that estimate the time delay only are described by Chan *et al.* (1980), (1981), Etter and Stearns (1981), Reed *et al.* (1981), Feintuch *et al.* (1981), Youn *et al.* (1982), (1983), David and Stearns (1983), Duttweiler (1983), Youn and Matthews (1984), Messer and Bar-Ness (1987), Ching and Chan (1988), Vasilev and Aidemirski (1990), Ho *et al.* (1990), (1992), (1993), Ching *et al.* (1991), Dokic and Clarkson (1992), Clarkson (1993), So and Ching (1993), So *et al.* (1994) and Ching and So (1994), amongst others.

2.2.5 Time delay estimation in the absence of other process parameters

In this section of the chapter, the estimation of ‘pure’ time delays is considered. Such time delays arise mainly in signal processing applications; in these applications, it is common to use the term ‘time difference of arrival’ (TDOA) rather than time delay. A number of classes of methods for estimating this parameter may be identified:

1. Methods for the estimation of a single time delay based on the cross-correlation of two signals.
2. Methods for multiple time delay estimation (in a MIMO environment) based on the cross-correlation of two signals.
3. Other time delay estimation methods.

The estimation of 'pure' time delays using gradient methods, considered in Section 2.2.4, are not reconsidered in this section.

The cross-correlation of two signals may be used to estimate the time delay between the two signals, as the time at which the cross-correlation term is maximised corresponds to the time delay estimate. Most methods of this type are off-line in nature; among the authors who discuss the cross-correlation method, and variations on the method, are Faure and Evans (1969), Knapp and Carter (1976), Cabot (1981), Carter (1981), Scarborough *et al.* (1981), Haas and Lindquist (1981), Hassab and Boucher (1981), Boucher and Hassab (1981), Stein (1981), Al-Hussaini and Kassam (1984), Azaria and Hertz (1984), Bradley and Kirlin (1984), Schwartzenbach and Gill (1984), Abatzoglou (1986), Fertner and Sjolund (1986), Hertz (1986), Al-Hussaini and El-Gayaar (1987), Gabr (1987), Weiss and Stein (1987), Krolik *et al.* (1988), Zheng and Feng (1988), Cusani (1989), Avitzour (1991), Kollar (1992), Gardner and Chen (1992a), (1992b), Gardner and Spooner (1992), Bar-Shalom *et al.* (1993), Clarkson (1993), Carter and Robinson (1993), Jacovitti and Scarano (1993), Kumar and Bar-Shalom (1993), Shen *et al.* (1993), Fong *et al.* (1994) and Izzo *et al.* (1994). Meyr and Spies (1984) estimate the time delay in closed loop using the method, and also define an algorithm for the tracking of a randomly varying time delay between two stochastic signals. During and Jansson (1993) define a variation on the cross-correlation algorithm that is suitable for an on-line implementation; the authors state that the estimation time of the algorithm implemented on a Texas Instruments TMS32020 signal processor is approximately 1 ms.

Other authors use the technique to estimate time delays in multi-input, multi-output environments or between multiple sensors and multiple targets. These algorithms are mostly off-line in nature, with examples of such algorithms described by Friedlander (1980), Ng and Bar-Shalom (1982), (1986), Tremblay *et al.* (1987) and Pallas and Jourdain (1991). Segal *et al.* (1991) and Antoniadis and Hero (1994) develop on-line, iterative algorithms for solving the multichannel time delay estimation problem.

Finally, other algorithms have been defined for the estimation of 'pure' time delays. One example of such an off-line algorithm is defined by Kenefic (1981), in which the time delay between two sensors may be found by determining the maximum of the probability density function (p.d.f.) of the delay from a given prior distribution. Nehorai and Morf (1982), Hertz and Reiss (1982), Azenkot and Gertner (1985), Chiu (1987), George and Goodman (1988), Jesus and Rix (1988), Moddemeijer (1989), Champagne *et al.* (1991), Jane *et al.* (1991), Yamada *et al.* (1991), Lourtie and Moura (1991), Boudreau and Kabal (1992), El-Hawary and Mbamalu (1993), Laguna *et al.* (1994), Manickan *et al.* (1994) and Koenig (1995) define other such off-line time delay estimation algorithms. Less attention appears to have been paid to the on-line implementation of non-cross correlation based algorithms, though one such algorithm is defined by Bethel and Rahikka (1987), who calculate recursively the p.d.f. of the time delay, from which an optimum estimate of the time delay may be determined. Algorithms based on the same approach are defined by Bethel and Rahikka (1990) and Bethel *et al.* (1995). Other on-line algorithms are defined by Namazi and Stuller (1987), Feder and Weinstein (1988), Namazi and Biswal (1992) and Blackowiak and Rajan (1995). The latter authors investigate the performance of a simulated annealing algorithm in the estimation of the amplitude scaling factors and the time delays of the separate arrivals in a signal composed of closely spaced arrivals with added noise. The method is particularly interesting as the cost function to be minimised has local minima that make the application of calculus based minimisation techniques (such as the Newton-Raphson gradient algorithm) difficult; the authors declare that the simulated annealing algorithm has the ability to slide through local minima.

O'Dwyer (1996a) discusses the algorithms outlined above that appear to merit further investigation; overall, however, the algorithms that estimate 'pure' time delays only are less useful, at least in control applications, than algorithms that facilitate the estimation of both the time delay and non-time delay model parameters.

2.3 Frequency domain methods for parameter and time delay estimation

In this section, both methods of estimating the frequency response of a process and methods for the subsequent estimating of the model parameters (including time delay) are considered, together with methods for the direct estimation of the model parameters based on the use of higher order spectra.

2.3.1 Frequency response estimation

The methods that have been defined for process frequency response estimation, in both open loop and closed loop environments, may be classified as follows:

1. The response to a sine wave input
2. The response to a pulse input
3. Correlation analysis
4. Spectral analysis
5. Methods based on the ratio of Fourier transforms
6. Optimisation methods
7. Cepstral analysis
8. The use of a relay in series with the process in closed loop and
9. Other methods.

Of course, the frequency range over which the model should be estimated needs to be defined. Generally, good frequency domain matching between the process and the model over a wide range of frequencies about the frequency where the phase lag of the process equals 180 degrees is desirable, particularly for controller design (Harris and Mellichamp (1980), Edgar *et al.* (1981), Wittenmark and Astrom (1984), Lee *et al.* (1990), Hang and Chin (1991) and Eskinat *et al.* (1993)).

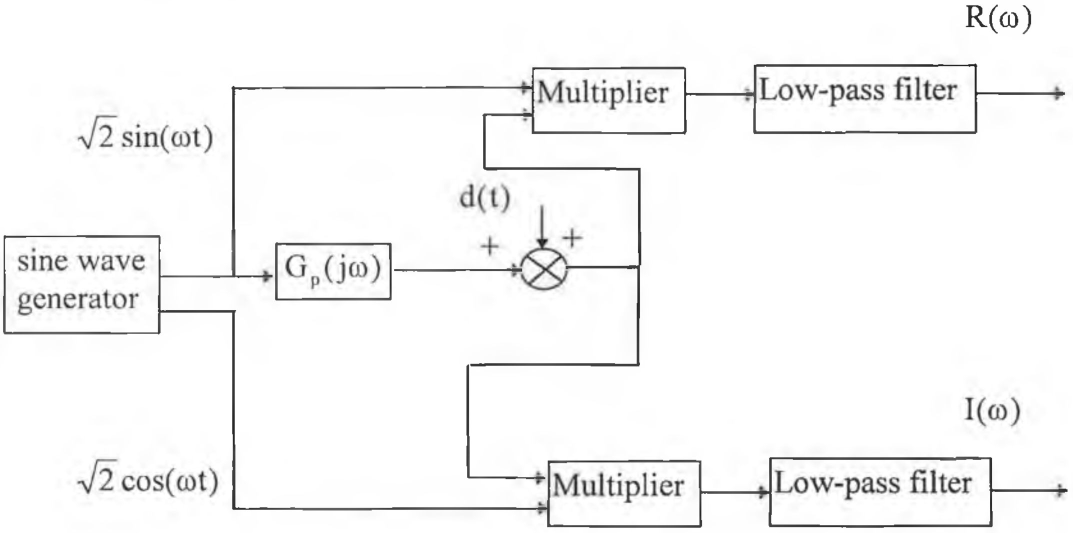
The frequency response of a process (in open loop) at any frequency may be determined by calculating the magnitude and phase of the process from its output when an appropriate sine wave is input to the process; however, the estimate obtained is sensitive to disturbances (Soderstrom and Stoica (1989), Larsen (1994)).

The frequency response of a process may also be found by determining the response of the process, in open loop, to a pulse input (Clements and Schnelle (1963)).

Good fitting of the magnitude response is found in experimental work carried out by these authors; the goodness of fit, however, does appear to worsen at higher frequency values. Other pulse response techniques are defined by Rajakumar and Krishnaswamy (1975), Harris and Mellichamp (1985), Seborg *et al.* (1989) and Smirthwaite *et al.* (1994).

The frequency response may be determined directly by correlation (Rake (1980), Unbehauen and Rao (1987), Soderstrom and Stoica (1989), Larsen (1994)). This approach may be represented in block diagram form, as shown in Figure 2.1.

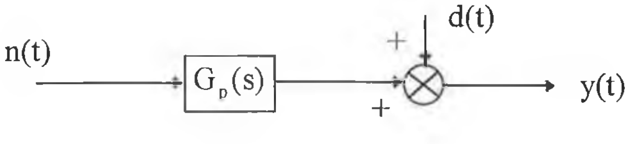
Figure 2.1: Correlation analysis (Unbehauen and Rao (1987))



The process frequency response at frequency ω, $G_p(j\omega)$, equals $R(\omega) + jI(\omega)$ (with d(t) being a disturbance). Larsen (1994) declares that the method is insensitive to step and white noise disturbances (due to the presence of the low-pass filters). However, long experiment times are often required to determine the process frequency response.

Spectral analysis techniques may also be used to calculate an estimate of the frequency response in both open loop and closed loop environments. In open loop, the process is represented as shown in Figure 2.2.

Figure 2.2: Open loop implementation



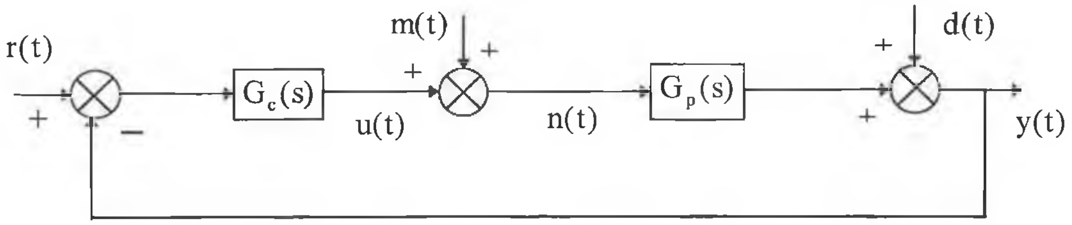
In this case, n(t) and d(t) are uncorrelated. The technique involves determining an estimate of the frequency response of the process, $G_p(j\omega)$, as follows:

$$G_p(j\omega) \approx S_{yn}(j\omega)/S_n(j\omega) \tag{2.11}$$

with $S_{yn}(j\omega)$ equal to the cross power spectral density of y(t) with respect to n(t) and $S_n(j\omega)$ equal to the power spectral density of n(t). The power spectral densities may be estimated using either the periodogram (sample spectrum) approach, which involves estimating the power spectral density in terms of the square of the corresponding discrete Fourier transform (Unbehauen and Rao (1987), Johannson (1993)) or the correllelogram approach, which involves estimating the relevant covariance functions, and calculating the estimates of the power spectral densities from the discrete Fourier transforms of these covariance functions (Unbehauen and Rao (1987)). Alternative methods defined by Schwartzenbach and Gill (1984) and Unbehauen and Rao (1987) may be used to estimate the phase response of the process, which is important for time delay estimation in particular. Chan *et al.* (1978), Hannan and Thomson (1981), Friedlander and Porat (1982), (1984), Chan and Miskowicz (1984) and Tachibana (1984) also use power spectral density techniques to calculate the model parameters and/or the time delay.

The closed loop system considered may be represented as shown in Figure 2.3.

Figure 2.3: Closed loop implementation



Wellstead (1986) shows that, if r(t), m(t) and d(t) are uncorrelated, then

$$G_p(j\omega) \approx S_{my}(j\omega)/S_{mn}(j\omega) \tag{2.12}$$

or

$$G_p(j\omega) \approx S_{ry}(j\omega)/S_{rn}(j\omega) \tag{2.13}$$

Approximations for the power spectral densities may be determined by using the discrete Fourier transform (DFT), for instance.

The frequency response of a process may also be obtained by using methods based on the ratio of Fourier transforms. In open loop (Figure 2.2), an estimate of $G_p(j\omega)$ may be expressed as

$$G_p(j\omega) \approx F[y(t)]/F[n(t)] \tag{2.14}$$

with F[ ] being the Fourier transform of the relevant signal. The Fourier transform terms may be approximated by using the DFT (when the resulting approximation is called the empirical transfer function estimate (ETFE)), by using the discrete time Fourier transform (DTFT) or by using a numerical integration method, such as the Adams-Moulton method. The applicability of such approximations is discussed in detail by Wellstead (1981), Ljung (1987), Unbehauen and Rao (1987), Johannson (1993) and Guillaume *et al.* (1996), among others. Other methods based on using Fourier transforms to estimate the time delay and/or the model parameters are defined by Hertz and Reiss (1982), Azenkot and Gertner (1985), Nagai (1986), Chiu (1987) and Boudreau and Kabal (1992).

In closed loop (Figure 2.3), and if r(t) and m(t) are uncorrelated, with F[d(t)] = 0, it will be proved in Chapter 4 that

$$G_p(j\omega) \approx F[y(t)]/F[n(t)] \tag{2.15}$$

As before, the Fourier transform terms may be approximated by using the DFT, the DTFT or an alternative numerical integration of the Fourier transform. The method for determining the frequency response in equation (2.15), when the Fourier transform terms are approximated by using DFT's, is also used by Lamaire *et al.* (1991) as a means of deriving a robust estimator of the process frequency response. Band-pass filters could be put on the input and output of the process so that F[d(t)] could be more reasonably assumed as zero, at one or more frequency values (Hagglund and Astrom

(1991), Ho *et al.* (1994)). A related possibility is to place a number of band-pass filters on the input and output of the process to determine the frequency response at a number of frequencies corresponding to the centre frequencies of the band-pass filters (Goberdhansingh *et al.* (1992)). Other authors that use Fourier transforms in closed loop as a means of estimating the frequency response of the process include Harris and Mellichamp (1980), Krishnaswamy *et al.* (1987), Koganezawa (1991), Hang and Sin (1991) and Hang *et al.* (1994b).

The frequency response of the process in open loop may be determined from the minimisation of a possibly multimodal cost function whose variables include either the DFT of the input and output signals to the process (Marshand and Fu (1985), Schoukens *et al.* (1988), Pintelon and Schoukens (1990), Pintelon and Van Biesen (1990), Kollar (1992)) or complex logarithmic frequency response data (Banos and Gomez (1995), Guillaume *et al.* (1995)). In closed loop, the maximum likelihood estimate of the process parameters may similarly be determined by the minimisation of a multimodal cost function whose variables include the DFT of the input and output signals to the process (Pintelon *et al.* (1992)). The input and output signals to the process are assumed to be correlated through a process noise term.

Hassab and Boucher (1976) estimate the time delay of a delayed and attenuated replica of a signal by the use of the natural logarithm of the magnitude squared of the output signal (called the power cepstrum of the signal). The authors state that when the technique is successful, the cepstrum yields a dominant peak away from the origin corresponding to the desired time delay. Barrett and Moir (1986) use cepstral methods for restoring the unknown phase-frequency information from the amplitude-frequency information that may be provided by the power spectral density techniques.

The relay autotuning method, developed first by Astrom and Hagglund (1984), may be used to determine one or more points on the frequency response of the process. The method involves the introduction of a relay element in parallel with the controller; the relay is switched in when process parameter estimation is required. The limit cycle provoked at the process output, as a result of the introduction of the relay element, may be analysed to determine approximations for the magnitude and frequency of the process at a process phase lag of 180 degrees. It is possible, as the authors suggest, to determine approximations for the magnitude and frequency of the process when the phase lag is 90 degrees, if an integrator is introduced in series with the relay. The authors also show that approximations for the magnitude and frequency may be obtained when the phase lag is between 90 and 180 degrees, by the introduction of appropriate hysteresis on the relay element. The method is developed further by Arzen (1987), Astrom and Hagglund (1988), Chiang and Yu (1993), Friman and Waller (1995) and Hwang (1995). Other related approaches using the relay autotuning method are proposed by Hagglund and Astrom (1989), (1991), Schei (1992), Astrom *et al.* (1993), Ho *et al.* (1994), Lundh and Astrom (1994), Lee *et al.* (1995b), Voda and Landau (1995b) and Shen *et al.* (1996a), (1996b), (1996c). In addition, the method may be applied to the estimation of the parameters of MIMO process models plus time delays, as detailed by Loh *et al.* (1993), Wu *et al.* (1994) and Friman and Waller (1994).

Other methods of estimating the frequency response of the process include estimating the magnitude and frequency of the process at a phase lag of 180 degrees (in closed loop), which is described by Balchen and Lie (1987); in this method, the system deviation signal is correlated with the excitation signal.

In conclusion, techniques that directly estimate the frequency response both in open loop and in closed loop have been well documented in the literature. The robustness of many of the techniques when closed loop identification is required, with process noise added to both the input and the output, is questionable; some authors address this problem by appropriate filtering of the process input and output signals prior to identification.

2.3.2 Parameter and time delay estimation using higher order spectra

2.3.2.1 Introduction

Higher order spectra (or polyspectra) are defined in terms of the higher order statistics (or cumulants) of a signal. The general motivations for the use of higher order spectral techniques are (1) to suppress additive, possibly coloured Gaussian noise that may be present on signals (2) to allow recovery of phase information from signals and (3) to detect and quantify nonlinearities in time series (Nikias and Petropulu (1993)). The use of higher order spectra is examined with special reference to the identification of the parameters of a SISO process model with a time delay, in both open loop and closed loop environments.

An important frequency domain approach to the identification of such a process model is to base the identification of the model parameters on the magnitude and phase response of the process. The use of second order statistics, which involve the calculation of the power spectral densities of the input and output signals to the process, gives rise to identifiability problems when both the input and output records are contaminated by even white and mutually uncorrelated noise sources (Delopoulos and Giannakis (1994)). However, because the higher order spectrum of (coloured) Gaussian signals is identically zero, adding coloured Gaussian noise of unknown spectrum to the process input or output does not affect the process frequency response estimation, if higher order spectral techniques are used.

The most common higher order spectra of a signal that are calculated are the third order spectrum (also called the bispectrum) and the fourth order spectrum (also called the trispectrum), as defined by Nikias and Petropulu (1993) and explored in detail by O'Dwyer (1996a). Cross-cumulants and the cross-bispectrum or cross-trispectrum may also be defined in a similar manner, using relevant process input and output signals (O'Dwyer (1996a)).

The bispectrum and trispectrum are special cases of the n^{th} order spectrum of a signal. Generally speaking, for computational reasons, the bispectrum of a signal is the most often calculated; the trispectrum of the signal may be calculated if the signal had zero (or very small) third order cumulants and larger fourth order cumulants. A symmetrically distributed random variable has a third order cumulant equal to zero, for instance (Mendel (1991)).

The cepstrum of higher order spectra may also be defined, as discussed by O'Dwyer (1996a).

2.3.2.2 Parameter estimation techniques using higher order spectra

It has been mentioned in Section 2.3.2.1 that an important means of determining the parameters of the process model in the frequency domain is to first determine the magnitude and phase variation of the process model with frequency. An intermediate stage may be to determine the bispectrum or trispectrum magnitude and phase estimates of the process. Nikias and Petropulu (1993) discuss a number of methods that have been defined for determining the bispectrum (or trispectrum) magnitude and phase estimates from the input and output data of a process; all of the methods involve the use of fast Fourier transforms. One of the difficulties about determining higher order spectra in this way is that only a finite set of data is used. Nikias and Petropulu (1993) and Matsouka and Ulrych (1984) also explain a number of the methods developed (by Bartlet *et al.* (1984) and Li and Ding (1994), amongst others) for the estimation of the process magnitude and phase (subsequently referred to as the Fourier magnitude and Fourier phase) from the bispectral magnitude and phase estimates of the process determined. Other papers in which details of these algorithms are provided include those by Haniff (1991), Matson (1991), Rangoussi and Giannakis (1991), Cheng and Venetsanopoulos (1992) and Li and Ding (1994). Pan and Nikias (1987) discuss the reconstruction of the Fourier phase from the corresponding trispectrum.

It is also possible to determine the bicepstrum and tricepstrum of the input and output data, as an intermediate stage to determining the Fourier gain and phase of the process. This is discussed by Alshebeili and Cetin (1990), Alshebeili *et al.* (1991) and Brooks and Nikias (1993).

The direct estimation of the process model parameters and the time delay using higher order spectral techniques (without first estimating the Fourier magnitude and Fourier phase of the process) does not appear to have been addressed in the literature. The estimation of the time delay between two signals (i.e. the estimation of a time delay term only, with no other dynamics considered) has been explored in detail by Nikias and Petropulu (1993), among others. The authors divide the methods used into the following categories:

(i) Conventional time delay estimation techniques based on third order statistics that involve estimating the time delay from the bispectral and cross-bispectral phases of the input and output signals to the process; Hinich and Wilson (1992), for example, estimate the time delay as the scaled difference between these phase estimates. Sato and Sasaki (1977), Sasaki *et al.* (1977), Nikias and Raghuveer (1987), Nikias and Pan (1988), Zhang and Raghuveer (1991) and Nikias and Mendel (1993) also outline methods of this type.

(ii) Parametric time delay estimation techniques, which involve modelling the time delay by a polynomial and estimating the polynomial coefficients; Nikias and Pan (1988), Tugnait (1991) and Delopoulos and Giannakis (1994) also outline these methods. In a more recent paper, Delopoulos and Giannakis (1996) extend the method of Delopoulos and Giannakis (1994) to the estimation of a process model (in rational polynomial form) in a closed loop environment, when both input and output data to the process is contaminated by additive noise having unknown cross-spectral characteristics.

(iii) Time delay estimation techniques based on the cepstrum of higher order spectra; Petropulu *et al.* (1988) and Reddy and Rao (1987) discuss such methods in detail.

(iv) Adaptive time delay estimation based on the parametric modelling of higher order cross-cumulants, which use third order cumulants and a gradient-like algorithm to estimate the time delay, where the additive noises on the signals are of spatially correlated Gaussian form with unknown correlation functions (Chiang and Nikias (1990)).

2.3.2.3 Conclusions

The following conclusions about the use of higher order spectral techniques for process parameter estimation may be drawn:

1. Conventional approaches for process frequency response estimation (based on the power spectrum, for instance) have a lower computational intensity and a requirement for a smaller number of data points than do the higher order spectral approaches. However, the higher order spectral approaches are robust to the presence of possibly mutually correlated, coloured Gaussian noise (or non-Gaussian noise, with a symmetric p.d.f.) added to both the process input and output.
2. The problem of process identification in closed loop using higher order spectra has not been completely resolved. The signals encountered in closed loop operation do not fit the requirement for the signals specified for process identification in all details; nevertheless, identification of the process parameters may be possible in certain situations in a closed loop environment (e.g. if a PRBS driving signal is added to the input of the process). Delopoulos and Giannakis (1996), in a recent paper, show that process identification in closed loop is possible using the third order cumulants of the process input and output signals.

It appears that the critical factor in the decision as to whether it is appropriate to use higher order spectra for process parameter estimation is the magnitude and nature of the additive noise present on both the input and output signals to the process. Johnson (1985) states that, realistically, noise terms either may have a known or estimated mean, covariance and distribution, or may have a constant bias component and a stochastic component having either a zero mean or being a filtered version of a white noise signal. For identification and control purposes in the self-tuning literature, the added noise is often considered to be modelled as the filtered version of a white noise signal. It is possible to reduce the effect of noise terms by pre-treatment of data before identification (e.g. if the noise term is drift on the input or output signals to the process, then the appropriate data could be filtered before identification); Ljung (1987) discusses a number of approaches in this area. In a closed loop process environment, there seems to be less justification for the use of higher order spectral techniques if a PRBS driving signal must be added at the process input, as such a signal will be uncorrelated to any noise signal and thus less computationally intense methods of process identification may be appropriate.

Overall, the use of higher order spectral techniques in system identification seems suited to a restrictive range of problems, in which noise signals on the input and output to the process cannot be effectively dealt with by pre-processing.

2.3.3 Model parameter estimation using frequency response data

The approaches to estimate the parameters of an appropriately ordered process model plus time delay, may be classified as follows:

1. Model parameter estimation using a graphical approach
2. Model parameter estimation using an analytical approach
3. Model parameter estimation using a least squares approach; the estimation of the parameters of a high order model plus time delay and the parameters of a low order model plus time delay will be considered separately and
4. Model parameter estimation based on relay identification.

The model parameter and time delay estimates may be determined graphically, from the Bode plots of the process; Deshpande and Ash (1983) and Seborg *et al.* (1989) apply the method to the estimation of the parameters of a FOLPD model and a SOSPD model, with Luyben (1983) fitting higher order transfer functions with time delay to the Bode plots. Seborg *et al.* (1989) identify the disadvantages of the method as the tediousness of the procedure, the introduction of errors in fitting parameters for second order models (and, by extension, higher order models) using a trial and error approach and that the method does not easily facilitate the identification of more general transfer function models, such as those with numerator dynamics.

The parameters may also be estimated analytically, from the frequency response of the process. Isermann *et al.* (1974), for instance, analytically determine the time constant and the time delay of a multiple pole process model. The model transfer function is

$$G_m(s) = \frac{K_m e^{-s\tau_m}}{(1 + sT_m)^n} \tag{2.16}$$

with K_m = model gain, T_m = model time constant and τ_m = model time delay. Then, the authors provide an estimate for the model time constant and time delay as follows:

$$T_m = \frac{1}{\omega}\sqrt{\left[\frac{K_m}{|G_p(j\omega)|}\right]^{\frac{2}{n}} - 1} \tag{2.17}$$

and

$$\tau_m = \frac{1}{\omega}\left[-\phi_p(j\omega) - n\tan^{-1}(\omega T_m)\right] \tag{2.18}$$

with $|G_p(j\omega)|$ = magnitude of the process transfer function and $\phi_p(j\omega)$ = phase of the process transfer function at a frequency ω. Special cases of these implementations are discussed by O'Dwyer (1992), (1993) and Hang *et al.* (1993), (1994b). Sundaresan and Krishnaswamy (1978), Koganezawa (1991) and O'Dwyer (1992) also consider other analytical methods of calculating the parameters of FOLPD and SOSPD models from the process frequency response.

Alternatively, the model parameters and time delay may be estimated by minimising the squared error between the process and the model in the frequency domain. For an arbitrary order model plus time delay, many of the techniques defined require the approximation of the time delay by an appropriate rational polynomial; the time delay as such is consequently not identified. Examples of such methods are discussed by Levy (1959), Whitfield (1986), (1987), Unbehauen and Rao (1987) and Hakvoort and Van den Hof (1994). Other authors, such as Dos Santos and De Carvalho (1990) explicitly estimate the parameters of an n^{th} order model plus time delay. These authors iteratively determine the estimates of the model order and the pole and zero values from the estimate of the time delay, with the estimate of the time delay calculated based on a least squares procedure using the phase plot. In a different method, Seborg *et al.* (1989) suggest that the parameters and the time delay of a process model could be estimated by selecting the value of the time delay iteratively and using the method of Levy (1959), for example, to determine the remaining process model parameters. Similar methods based on this multiple model estimation technique have been well explored in the time domain. However, the multiple model estimation method is computationally intensive. In a more recent paper, Young *et al.* (1995) estimate the model parameters and time delay of a linear process using a recursive non-linear estimation technique in the frequency domain. The authors mention that it is possible that the parameters (and delay) identified may correspond to a local minimum of the cost function used, rather than a global minimum.

It is also possible to fit a low order model plus delay to the process response, in a least squares sense. Lilja (1988), for instance, calculates a FOLPD model of a high order process from four points on the frequency response of the process. The non-time delay parameters are determined by minimising an appropriate quadratic cost function; the time delay is determined separately by minimising a multimodal cost function using a modified Newton-Raphson algorithm, though convergence of the time delay estimate to the correct value of the time delay is consequently not guaranteed. Nevertheless, the author gives advice on strategies to determine the best estimate of the time delay. A simulation result provided by the author shows that a process of order 16 is well approximated by a FOLPD model, in a frequency range corresponding to a phase lag range of 0 to 180 degrees. Other authors that describe algorithms of this type include Seborg *et al.* (1989) and Palmor and Blau (1994).

Finally, the parameters of the process model may also be identified by analysing the process output when a relay is switched into the closed loop compensated system in place of the controller. It is possible to approximate the limit cycle output as a sinusoid (this is the basis of the approach of Astrom and Hagglund (1984) for controller tuning). However, it is also possible to analyse the limit cycle output without any such approximation being taken. Lee and Sung (1993), for instance, calculate the time constant and time delay of a FOLPD model using this approach, as follows:

$$T_m = \frac{0.5t_0}{\ln[\frac{1+(a/K_m d)}{1-(a/K_m d)}]} \qquad (2.19)$$

and

$$\tau_m = T_m \ln[\frac{1+e^{t_0/2T_m}}{2}] \qquad (2.20)$$

with t_0 = period of oscillation of the limit cycle, a = amplitude of the oscillation and d = relay amplitude (K_m would need to be known *a priori*). Arzen (1987), Li *et al.* (1991), Chang *et al.* (1992), Leva (1993), Benouarts and Atherton (1994) and Egan (1994) also describe algorithms of this type. Indeed, it appears reasonable that further work in this area is possible, as many authors use such relay compensator techniques for autotuning rather than for model parameter estimation; to this end, a development of the method defined by Egan (1994) is outlined briefly in Chapter 8 and is discussed in detail by O'Dwyer (1996k).

In conclusion, the approaches for process model parameter and time delay estimation that have been proposed in the literature in the frequency domain have been briefly documented above. The range of frequency values over which the process is estimated can be specified, depending on the application.

2.4 Other methods of process parameter and/or time delay estimation

The previous sections of the chapter have described in detail well-defined methods of process parameter and time delay estimation. Other methods do not easily fall into the categories discussed earlier; in this section, parameter and/or time delay estimation using neural networks, using process order identification methods, using the delta operator and using genetic algorithms are outlined.

Neural networks may be used for the identification and control of non-linear processes (Narendra and Parthasarathy (1990)). The identification and control of time delay processes using neural networks is a subject of recent research. Bhat and McAvoy (1992), for instance, propose a detailed method to strip a back propagation neural network (BPN) to its essential weights and nodes to give it its simplest possible structure; the authors show that the stripping algorithm is capable of identifying the time delay and order of a FOLPD process (in the discrete time domain). Other authors that discuss the identification and control of processes using neural networks include Megan and Cooper (1992), who present a neural network approach to adaptive control by analysing the relationship between the error pattern and the corresponding adjustment needed in the gain and time constant of a first order lag (FOL) model of a process, and Hinde and Cooper (1994), (1995) who explore the use of a passive adaptive algorithm which updates the process model and the controller in closed loop by taking advantage of naturally occurring dynamic events, rather than injecting perturbations into the system to create dynamic events. Cheng *et al.* (1995) identify a non-linear dynamic process with unknown and possibly variable time delay using an internal recurrent neural network. However, it is true to say that the use of neural networks for the identification of processes with time delays is in its infancy.

Process order estimation strategies may also be used to estimate the time delay of a process (in the discrete time domain), since the time delay appears as an increase in the model order of the numerator transfer function. Process order identification strategies for SISO systems have been described by Unbehauen and Gohring (1974), Van den Boom and Van den Enden (1974), Wellstead (1978), Stoica *et al.* (1986), Unbehauen and Rao (1987), Soderstrom and Stoica (1989), Niu *et al.* (1990), O'Donnell (1991) and Liang *et al.* (1993). Process order identification strategies for

MIMO systems have been defined by Guidorzi (1975), (1981), Tse and Weinert (1975), Lin and Wu (1982), Van Overbeek and Ljung (1982), Stoica (1983), Zhang *et al.* (1985), Li (1985), Chen and Guo (1987), Guo *et al.* (1989), Chen and Zhang (1990), Zhang and Chen (1990), Guillaume *et al.* (1992), Gu and Misra (1992), Glentis and Kalouptsidis (1992) and Niu and Fisher (1994). The estimation of the time delay using these strategies would, however, be conditioned on the order of the non-delay part of the process being known *a priori*.

It is also possible to estimate the process parameters using the delta operator rather than the z (or shift) operator. The delta operator (also known as the Euler operator) is defined as follows:

$$\delta = \frac{z-1}{T_s} \tag{2.21}$$

where T_s equals the sample time. Wellstead and Zarrop (1991) show that the region of stability for z domain poles (i.e. the unit circle) translates into a circle of centre $(-1/T_s, 0)$ and radius $1/T_s$ in the delta domain. Thus, as the sampling rate is increased, the stability region defined in the delta domain approaches that of the continuous time domain. The following advantages are claimed for the delta operator representation over the shift operator representation:

(a) The representation of discrete systems in the delta domain avoids the problems of coefficient sensitivity in recursive digital filters at high sample frequencies, seen in the z domain (Goodall (1990)).

(b) A related advantage is that the delta operator allows superior finite word length coefficient representation (Middleton and Goodwin (1986), (1990)) under the assumption that the sample time is chosen according to the normally quoted rules of thumb.

(c) A further advantage of the delta operator is that it "almost always" has less roundoff noise associated with it than does the corresponding z operator (Middleton and Goodwin (1986), (1990)).

(d) Middleton and Goodwin (1986), (1990) and Goodwin *et al.* (1988), (1992) declare that for parameter estimation, the least squares solving of a set of linear equations is better conditioned for models represented by the delta transform.

Terrett and Downing (1993), (1994) use the delta operator for system identification on a fixed point DSP using the RLS algorithm. The authors show that the delta operator is numerically more robust than the shift operator, at the cost of an additional computational requirement (this point is also made by Goodwin *et al.* (1992)). Other authors that discuss the use of the delta operator for system identification purposes include Goodwin *et al.* (1988), Middleton and Goodwin (1990), Wilkinson *et al.* (1991) and Jabbari (1991). The use of the delta domain for the estimation of the time delay as well as the model parameters has not been considered in the literature, with the exception of the method defined by Keviczky and Banyasz (1992), who identify the time delay index as the sum of the product of each numerator parameter identified and its index, divided by the sum of the identified numerator parameters, where both sums are taken up to a defined maximum time delay index. The process is modelled as a SOSPD model, with identification of the time delay taking place in the delta domain. There does seem to be scope to estimate the time delay (and the other model parameters) in the delta domain, using techniques similar to those used in the z domain (e.g. the overparameterisation of the process model); the method of Keviczky and Banyasz (1992), for instance, is an analogue of a method defined by these authors (Keviczky and Banyasz (1988)) in the z domain. The use of the delta domain for system identification is also explored by O'Dwyer (1996a), (1996l).

Finally, the use of genetic algorithms for process identification (including time delay identification) is beginning to attract interest. Genetic algorithms search from a population of points, use information about the cost function (rather than its derivative or other auxiliary knowledge used by gradient algorithms) and have a random component (quantified as a mutation rate) that helps drive the model parameters towards values corresponding to the global minimum of the appropriate cost function. These algorithms tend to be very computationally intensive; Kristinsson and Dumont (1992) declare, for instance, that a genetic algorithm is perhaps fifty times more computationally intensive than is a recursive instrumental variable (RIV) system identification algorithm. Genetic algorithms are considered to be one extreme solution to the exploitation-exploration trade-off, as described by Renders and Flasse (1996); genetic algorithms trade-off large computation time, and poor accuracy of the global minimum, with reliability in calculating the global minimum. The authors consider the use of gradient algorithms to be another solution to the exploitation-exploration trade-off; other solutions, such as the use of multiple model estimation methods, are of course also possible. In an interesting recent application, Yang *et al.* (1996) use a genetic algorithm to estimate the denominator parameters and time delay of a reduced order process model, while using the less computationally intensive least squares algorithm to subsequently determine the numerator parameters (which is a linear problem).

2.5 Conclusions

This chapter has considered a wide variety of methods for time delay and model parameter estimation, in both the continuous time and discrete time domains. The methods have been discussed in detail by O'Dwyer (1996a), in which comparisons have been drawn where appropriate between methods. The wide spectrum of methods covered means that an overall conclusion as to the best method to use is not appropriate. However, it is possible to indicate the areas and methods in which original work may be done.

1. A major section of the chapter has been devoted to the use of gradient methods for model parameter and time delay estimation. It has been decided to investigate fully the methods defined by Durbin (1984a), (1984b), (1985)), which facilitate identification of the model parameters and the time delay in open loop, because of the potential of the methods to estimate the parameters quickly, even in the presence of bias and noise terms. Alternative polynomial approximations to the time delay than those taken by the author will also be considered. This work is carried out in Chapter 3. In addition, it has been decided to investigate closed loop methods for estimating the model parameters and time delay, based on the work done by Marshall (1979), (1980) and Bahill (1983). Alternative updating algorithms for the parameters will be defined, considering fewer assumptions on system behaviour than are considered by the above authors; in addition, the parameter updating strategies will be extended to the estimation of the parameters of higher order models than the strategies considered previously. This work is reported in Chapter 7.

2. The frequency domain is a very natural domain for time delay estimation, as was mentioned in Section 2.1. It has been decided to estimate the process frequency response, both in open loop and in closed loop, using approaches based on the ratio of Fourier transforms, and the ratio of appropriate power spectral densities. It has also been decided to investigate a method that combines an analytical approach that gives initial estimates of the parameter and time delay values, with a least squares approach using a gradient algorithm that updates these estimates to more accurate model parameter estimates. Both of these topics are discussed in Chapter 4.

3. Other areas in which original work may be done are

(a) The estimation of the process parameters (including time delay) based on an analytical description of the actual process output, rather than on parameters calculated assuming a sinusoidal process output, when a relay is introduced in series with the process in closed loop. This work is detailed by O'Dwyer (1996k).

(b) The use of relevant discrete time algorithms for the estimation of the parameters plus time delay of MIMO process models; in particular, the application of overparameterisation methods, such as the algorithm defined by Wong and Bayoumi (1982). Some preliminary work in this area is detailed by O'Dwyer (1996l).

CHAPTER 3

Open loop time domain gradient methods of parameter and time delay estimation

3.1 Introduction

Gradient methods of parameter estimation are based on updating the parameter vector using a vector that depends on information about the cost function to be minimised (which is equal to the sum of the squared error between the process and model outputs). The use of gradient algorithms for model parameter and time delay estimation is discussed in detail in Chapter 2.

The review and analysis of the available literature has revealed the close relationship between many of the methods used for time delay estimation using gradient methods. For control applications, the estimation of non-time delay parameters as well as the time delay is frequently required (e.g. for compensator design). Therefore, it has been decided to concentrate on methods that intrinsically estimate both model parameters and time delay. The best model to use for identification purposes is a vexed question as it depends, amongst other factors, on the data quality available (see Chapter 1); a cautious approach, which has been implemented in this chapter, is to identify the parameters of a FOLPD model (Newell and Lee (1989)). It was decided to investigate fully the method defined by Durbin (1984a), (1984b), (1985) because of its potential to estimate the parameters quickly, even in the presence of bias inputs and noise terms. In this method, the process is assumed to be modelled by a FOLPD model. The procedures developed may be applied to the estimation of the parameters of higher order models of a general order process, though the complexity of the development is greatly increased; this application will be discussed further in the conclusions of the chapter. The gradient algorithms are implemented by determining the partial derivatives of the error between the process output and the discretized model output, with respect to the gain, time constant and time delay. These partial derivatives are subsequently used to update the model parameters. However, prior to calculating the partial derivative of the error with respect to the time delay, the time delay variation (which equals the process time delay minus the model time delay) is approximated by a rational polynomial. Such an approximation will be valid for small values of the time delay variation; the approach is appropriate, as the use of the Gauss-Newton gradient algorithm, for example, depends on the difference between the estimated parameters and the actual parameters being small (as the Gauss-Newton algorithm is derived from a second order Taylor's series expansion of the cost function about the optimum parameter vector). The most appropriate rational polynomial to use may be determined by finding the relationship between the mean squared error (MSE) function between the process and model outputs and the time delay; this relationship, which may be determined both analytically and in simulation, must be unimodal about the process time delay for successful application of the gradient descent algorithms, as must the corresponding relationship of the MSE function to the process gain and time constant values.

The chapter considers, both analytically and in simulation, the convergence of the parameters of a FOLPD model to corresponding process parameters. The convergence of the non-time delay model parameters is considered first, when the time delay is assumed known. Subsequently, the convergence of the model time delay is discussed, when the non-delay model and process parameters are identical. A number of theorems are developed; one theorem considers convergence of the model time delay index in the idealised case, when it is assumed that the process time delay is an integer multiple of the sample period, and is known *a priori*. Subsequent work assumes, more realistically, that the process time delay is unknown *a priori*; convergence is considered when the process time delay is an integer multiple of the sample period, and a real multiple of the sample period, as well as when the previous model output is used to calculate the new model output. Finally, the convergence of all of the model parameters is considered, when all of the process parameters are unknown, in the idealised and realistic cases mentioned above. This structure allows a comprehensive exploration of the issues.

The author has considered the use of four gradient algorithms: the Levenberg-Marquardt algorithm, the Gauss-Newton algorithm, the steepest descent algorithm (Ljung (1987)) and the least mean square (LMS) algorithm (Widrow and Stearns (1985)). The implementation of these algorithms is described in Chapter 2. O'Dwyer and Ringwood (1994a), (1994b) show the estimation of the parameters of a FOLPD model using these algorithms, with each algorithm facilitating the updating of the parameters in a broadly similar manner (at least for the simulations taken). The simulation results quoted in this chapter will use the Levenberg-Marquardt gradient algorithm for updating the parameters, with no loss of generality, as the procedures developed to facilitate convergence of the model parameters to the process parameters are appropriate for the use of any of the gradient algorithms mentioned.

3.2 Rational polynomial approximation of the time delay variation

It has been stated in Section 3.1 that prior to calculating the partial derivative of the error with respect to the time delay, the time delay variation, r, is approximated by a rational polynomial. The two first order approximations to the time delay variation considered are as follows:

$$\text{Taylor:} \qquad e^{-sr} \approx 1 - sr \tag{3.1}$$

$$\text{Pade:} \qquad e^{-sr} \approx \frac{1 - 0.5sr}{1 + 0.5sr} \tag{3.2}$$

The MSE function between the process and model outputs was calculated analytically, when the time delay variation was represented by each of these approximations in turn; it is assumed that the process time delay is an integer multiple of the sample period. It was determined that the MSE performance surface was unimodal with respect to the model time delay index when the first order Taylor's series approximation was used. These calculations are done in the discrete time domain, as integer values of the process time delay appear as appropriate power terms on the numerator transfer function of the process, in this domain; in addition, a standard procedure has been defined to calculate the MSE surface, by Widrow and Stearns (1985), in the domain. These calculations are performed in subsequent sections of this chapter. The use of the first order Pade approximation produced a non-quadratic MSE performance surface, which is non-unimodal in the model time delay index. This development is given by O'Dwyer (1996e). The relationship between the MSE function and the model gain and time constant terms separately is unimodal; no approximation is used for the time delay variation when these parameters are being updated.

The use of higher order approximations for the time delay variation is possible; some of the second order approximations that may be used in these circumstances are as follows:

$$\text{Taylor: } e^{-sr} \approx 1 - sr + 0.5s^2r^2 \tag{3.3}$$

$$\text{Pade: } e^{-sr} \approx \frac{1 - 0.5sr + 0.0833s^2r^2}{1 + 0.5sr + 0.0833s^2r^2} \tag{3.4}$$

$$\text{Marshall (1979) : } e^{-sr} \approx \frac{1 - 0.0625s^2r^2}{1 + 0.0625s^2r^2} \tag{3.5}$$

$$\text{Product (Piche (1990)): } e^{-sr} \approx \frac{1 - 0.5sr + 0.125s^2r^2}{1 + 0.5sr + 0.125s^2r^2} \tag{3.6}$$

$$\text{Laguerre (Piche (1990)): } e^{-sr} \approx \frac{1 - 0.5sr + 0.0625s^2r^2}{1 + 0.5sr + 0.0625s^2r^2} \tag{3.7}$$

Paynter (Robinson and Soudack (1970)):

$$e^{-sr} \approx \frac{1}{1 + sr + 0.4054s^2r^2} \tag{3.8}$$

Product (Gradshteyn and Ryzhik (1980)):

$$e^{-sr} \approx \frac{1 - 0.5sr + 0.1013s^2r^2}{1 + 0.5sr + 0.1013s^2r^2} \tag{3.9}$$

Direct Frequency Response (Stahl and Hippe (1987)):

$$e^{-sr} \approx \frac{1 - 0.49sr + 0.0954s^2r^2}{1 + 0.49sr + 0.0954s^2r^2} \tag{3.10}$$

There are an infinite number of higher order approximations that may be used for the time delay variation; just one of these approximations is the third order approximation defined by Marshall (1979):

$$e^{-sr} \approx \frac{1 - (0.167sr)^3}{1 + (0.167sr)^3} \tag{3.11}$$

The use of the second order Taylor's series approximation depends on the use of a higher order model for the process than a FOLPD model; the other approximations may be used with a FOLPD process model. It is shown by O'Dwyer (1996e) that when the MSE surface was calculated versus model time delay index, over a large number of samples, for the approximations in equations (3.3) to (3.11), unimodality was achieved only when a second order Taylor's series approximation was used for the time delay variation (when the process is modelled by a second order lag plus time delay model).

Two examples of a non-unimodal MSE surface are provided below, when a second order Pade approximation and a second order Product approximation (as defined by Piche (1990)) is used for the time delay variation (Figures 3.1 and 3.2, respectively); the point 'x' marks where r = 0 (or $g_p = g_m$) in each case.

Figure 3.1: MSE surface (Pade)

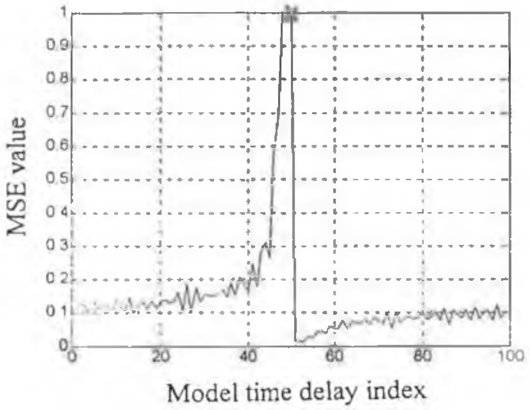
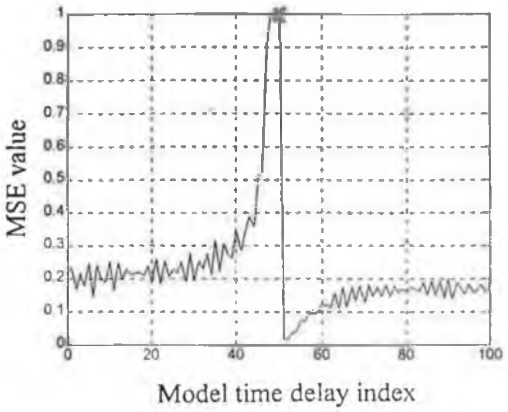


Figure 3.2: MSE surface (Product)



The condition that only the use of either a first order Taylor's series approximation or a second order Taylor's series approximation for the time delay variation will guarantee unimodality of the resulting MSE function versus model time delay is related to the z domain models calculated (using the zero order hold equivalence approach) when various approximations are used (O'Dwyer (1996e)). The poles of the z domain model are always within the unit circle when either a first order Taylor's series approximation or a second order Taylor's series approximation for the time delay variation is used, but one or more poles are either on or outside the unit circle when any other approximation is used, for at least some values of the model time delay index. It is perhaps not surprising that the resultant generation of an unstable discrete domain model does not facilitate convergence of the model time delay index to the process time delay index. Even in cases where unimodality of the MSE surface is achieved over a large range of model time delay index values, an infinite spike always exists on the MSE surface when r = 0 for all approximations except the Taylor's series approximations taken (e.g. Figures 3.1, 3.2). Thus, an exact estimate of the process time delay, using the gradient method, will not be possible in these circumstances. Unfortunately, Theorem 3.1 (Section 3.3) proves that satisfactory values of the non-time delay model parameters will not be estimated unless an exact estimate of the process time delay is determined.

3.3 Convergence of the non-delay model parameters

This section deals with the convergence of the non-delay model parameter estimates to the non-delay process parameter estimates using gradient methods; it is desired to prove that when the model time delay index equals the process time delay index, then the gradient algorithms may provide successful convergence of the model gain and time constant values to the process gain and time constant values, respectively.

Theorem 3.1: For a first order discrete stable system, the MSE performance surface is minimised when the model gain equals the process gain and the model time constant equals the process time constant, under the following conditions:
(a) The model time delay index equals the process time delay index
(b) Measurement noise is assumed uncorrelated with the process input and output and
(c) The input to the process and the model is assumed to be a white noise input.
Proof: The process difference equation is

$$y_1(n) = e^{-T_s/T_p} y_1(n-1) + K_p(1 - e^{-T_s/T_p}) u(n - g_p - 1) + w(n) \tag{3.12}$$

with T_p = process time constant, K_p = process gain and $\tau_p = g_p T_s$, T_s = sample period, g_p = process time delay index; w(n) = measurement noise.

The model difference equation is (assuming the previous process output is used in its calculation)

$$y_{m1}(n) = e^{-T_s/T_m} y_1(n-1) + K_m(1 - e^{-T_s/T_m}) u(n - g_m - 1) \tag{3.13}$$

with K_m = model gain, T_m = model time constant and g_m = model time delay index. The difference between the process and model output, $e_1(n)$, is (from equations (3.12) and (3.13))

$$\begin{aligned} e_1(n) = y_1(n) - y_{m1}(n) &= (e^{-T_s/T_p} - e^{-T_s/T_m}) y_1(n-1) \\ &+ K_p(1 - e^{-T_s/T_p}) u(n - g_p - 1) - K_m(1 - e^{-T_s/T_m}) u(n - g_m - 1) + w(n) \end{aligned} \tag{3.14}$$

The procedure defined by Widrow and Stearns (1985) may be used to calculate the MSE performance surface as follows:

$$E[e_l^2(n)] = r_{y_l y_l}(0) + r_{ww}(0) + \frac{1}{2\pi j}\oint [G_{ml}(z^{-1})\Phi_{uu}(z)G_{ml}(z) - 2\Phi_{y_l u}(z)G_{ml}(z)]\frac{dz}{z} \qquad (3.15)$$

with $\Phi_{uu}(z) = \sum_{n=-\infty}^{\infty} r_{uu}(n)z^{-n}$, $\Phi_{y_l u}(z) = \sum_{n=-\infty}^{\infty} r_{y_l u}(n)z^{-n}$, $r_{y_l y_l}(n)$, $r_{uu}(n)$ and $r_{ww}(n)$ being the autocorrelation functions of $y_l(n)$, $u(n)$ and $w(n)$ respectively; $r_{y_l u}(n)$ is the cross-correlation of $y_l(n)$ and u(n). The model $G_{ml}(z)$ corresponds to the output $y_{ml}(n)$.

Using the residue theorem to calculate the closed curve integral, the MSE function is calculated (from equation (3.15)) to be (O'Dwyer (1996m)):

$$E[e_l^2(n)] = \frac{K_p^2(1-e^{-T_s/T_p})^2}{(1-e^{-2T_s/T_p})} + \frac{K_m^2(1-e^{-T_s/T_m})^2}{(1-e^{-2T_s/T_m})} - \frac{2K_pK_m(1-e^{-T_s/T_p})(1-e^{-T_s/T_m})}{(1-e^{-T_s/T_p}e^{-T_s/T_m})}e^{-T_s(g_p-g_m)/T_m} + r_{ww}(0) \qquad (3.16)$$

The MSE function is minimised when $\partial E[e_l^2(n)]/\partial K_m$ and $\partial E[e_l^2(n)]/\partial(1/T_m)$ equal zero simultaneously. The required calculations, determined by partially differentiating equation (3.16) (O'Dwyer (1996m)), show that (assuming $g_p = g_m$)

$$\frac{\partial E[e_l^2(n)]}{\partial K_m} = 0 \Rightarrow K_m = \frac{K_p(1-e^{-T_s/T_p})(1-e^{-2T_s/T_m})}{(1-e^{-T_s/T_m})(1-e^{-T_s/T_p}e^{-T_s/T_m})} \qquad (3.17)$$

and

$$\frac{\partial E[e_l^2(n)]}{\partial(1/T_m)} = 0 \Rightarrow K_m = \frac{K_p(1-e^{-T_s/T_p})^2(1-e^{-2T_s/T_m})^2}{(1-e^{-T_s/T_m})^2(1-e^{-T_s/T_p}e^{-T_s/T_m})^2} \qquad (3.18)$$

If both $\partial E[e_l^2(n)]/\partial K_m$ and $\partial E[e_l^2(n)]/\partial(1/T_m)$ equal zero simultaneously, then it may be deduced from equations (3.17) and (3.18) that $T_m = T_p$ and $K_m = K_p$. □

A corollary to this theorem is that if $g_p \neq g_m$, then the MSE function is not minimised when $K_m = K_p$ and $T_m = T_p$. This means that the model time delay index must converge to the process time delay index before convergence of the model gain and time constant values to the process gain and time constant values, respectively, is possible. A further corollary to this theorem is that the MSE function is not minimised when $K_m = K_p$ unless $g_m = g_p$ and $T_m = T_p$, and the MSE function is not minimised when $T_m = T_p$ unless $g_m = g_p$ and $K_m = K_p$. These conclusions have been constant features of the simulation results taken for model parameter and time delay updating using gradient methods.

The unimodality of the MSE function with respect to the parameters individually was demonstrated, in typical simulation results, by plotting the MSE function versus the corresponding parameter, as shown in Figures 3.3 and 3.4. For Figure 3.3, $T_p = T_m$ and $g_p = g_m$ with $K_p = 100$ and for Figure 3.4, $K_p = K_m$ and $g_p = g_m$ with $T_p = 100$; for both of these plots, the MSE is calculated based on equation (3.16), with measurement noise assumed absent.

Figure 3.3: MSE vs. Model gain

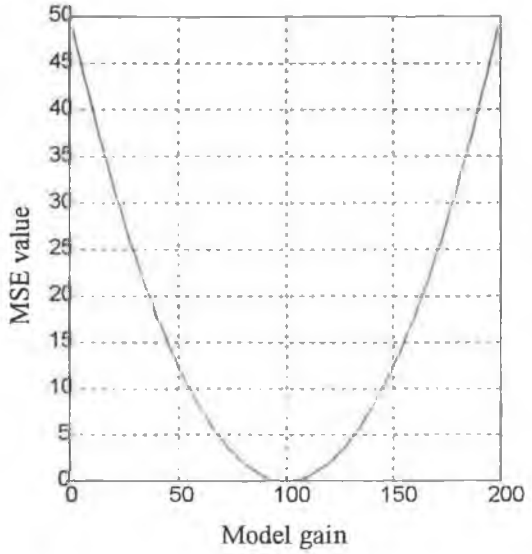
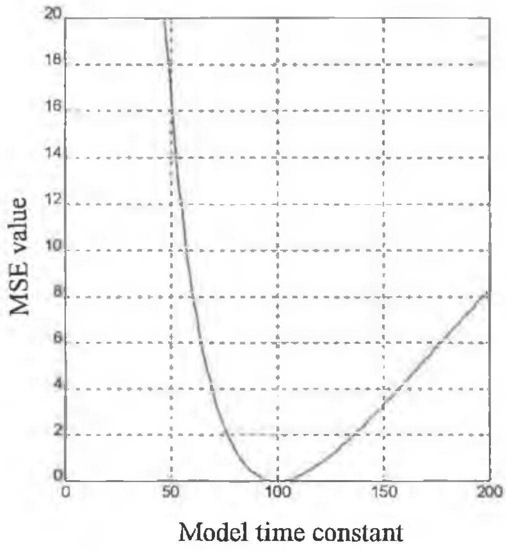


Figure 3.4: MSE vs. Model time constant



These plots confirm that the MSE function is quadratic with respect to the model gain, though it is not quadratic with respect to the model time constant (equation (3.16)).

3.4 Convergence of the model time delay

This section of the chapter will consider the use of the gradient algorithm for updating the time delay only, with the non-time delay process and model parameters put equal. The gradient algorithm used, depending as it does on the partial derivative of the cost function with respect to the parameter value (equation (2.4)), will be a function of the error between the process and the output, and the partial derivative of the error between the process and the output with respect to the parameter value. The cases outlined below are considered; these cases are chosen to comprehensively cover the implementations possible.

1. The error and the partial derivative of the error with respect to the time delay variation are calculated by using a first order Taylor's series approximation for the time delay variation. This will be referred to as Case 1 in subsequent work in this section of the chapter. This is an idealised implementation, as the process time delay is assumed known *a priori* (and is assumed to be an integer multiple of the sample period).
2. The partial derivative of the error with respect to the time delay variation is calculated by using a first order Taylor's series approximation for the time delay variation and the error is calculated based on using a FOLPD process model. In this case, updating of the model time delay when it is both an integer multiple of the sample period, and a real multiple of the sample period, is considered. This will be referred to as Case 2 in subsequent work in this section of the chapter. This case provides a more realistic implementation than Case 1 above, as the process time delay is not assumed known *a priori*.
3. The partial derivative of the error with respect to the time delay variation is calculated by using a first order Taylor's series approximation for the time delay variation and the error is calculated based on using a FOLPD process model. The previous model output is used to calculate the new model output. This will be referred to as Case 3 in subsequent work in this section of the chapter; as in Case 2, the time delay is not assumed known *a priori*, though it is assumed that the process time delay is an integer multiple of the sample period.

3.4.1 Convergence of the model time delay - Case 1

<u>Theorem 3.2</u>: For a first order discrete stable system of known gain and time constant, the MSE performance surface versus model time delay index is unimodal with a minimum value of the MSE occurring when the model time delay index equals the process time delay index, under the following conditions:

(a) The time delay variation is approximated by a first order Taylor's series approximation

(b) The measurement noise is uncorrelated with the process input

(c) The resolution on the process time delay is assumed to be equal to one sample period and

(d) The error and the partial derivative of the error with respect to the time delay variation are calculated based on using the first order Taylor's series approximation for the time delay variation.

<u>Proof</u>: The process difference equation, $y_2(n)$, is

$$y_2(n) = e^{-T_s/T} y_2(n-1) + K(1 - e^{-T_s/T}) u(n - g_p - 1) + w(n) \tag{3.19}$$

with $T_p = T_m = T$, $K_p = K_m = K$. The corresponding model difference equation, calculated by substituting a first order Taylor's series approximation for the time delay variation, is (assuming the previous process output is used in its calculation)

$$y_{m2}(n) = e^{-T_s/T} y_2(n-1) - \frac{K(g_p - g_m)T_s}{T} u(n - g_m) - K(e^{-T_s/T} - 1 - \frac{(g_p - g_m)T_s}{T}) u(n - g_m - 1) \tag{3.20}$$

Therefore, from equations (3.19) and (3.20), $e_2(n) = y_2(n) - y_{m2}(n)$ is given by

$$\begin{aligned} e_2(n) = K[(1 - e^{-T_s/T}) u(n - g_p - 1) &+ \frac{(g_p - g_m)T_s}{T} u(n - g_m) \\ &+ (e^{-T_s/T} - 1 - \frac{(g_p - g_m)T_s}{T}) u(n - g_m - 1)] + w(n) \end{aligned} \tag{3.21}$$

The MSE performance surface, $E[e_2^2(n)]$, may then be calculated as

$$\begin{aligned}
E[e_2^{\,2}(n)] = {} & E[K^2(1-e^{-T_s/T})^2u^2(n-g_p-1)+\frac{K^2(g_p-g_m)^2T_s^2}{T^2}u^2(n-g_m)] \\
& +E[\frac{2K^2(g_p-g_m)T_s}{T}(1-e^{-T_s/T})u(n-g_m)u(n-g_m-1)] \\
& +E[K^2(e^{-T_s/T}-1-\frac{(g_p-g_m)T_s}{T})^2u^2(n-g_m-1)] \\
& +E[2K(e^{-T_s/T}-1-\frac{(g_p-g_m)T_s}{T})u(n-g_m-1)w(n)+w^2(n)] \\
& +E[2K^2(1-e^{-T_s/T})(e^{-T_s/T}-1-\frac{(g_p-g_m)T_s}{T})u(n-g_p-1)u(n-g_m-1)] \\
& +E[\frac{2K^2(g_p-g_m)T_s}{T}(e^{-T_s/T}-1-\frac{(g_p-g_m)T_s}{T})u(n-g_m)u(n-g_m-1)] \\
& +E[2K(1-e^{-T_s/T})u(n-g_p-1)w(n)+\frac{2K(g_p-g_m)T_s}{T}u(n-g_m)w(n)] \qquad (3.22)
\end{aligned}$$

Therefore, it may be shown that (O'Dwyer (1996m))

$$\begin{aligned}
E[e_2^{\,2}(n)] = {} & 2K^2(1-e^{-T_s/T})^2[r_{uu}(0)-r_{uu}(g_p-g_m)]+\frac{2K^2(g_p-g_m)^2T_s^2}{T^2}[r_{uu}(0)-r_{uu}(1)] \\
& +\frac{2K^2(1-e^{-T_s/T})T_s}{T}(g_p-g_m)[r_{uu}(0)-r_{uu}(1)+r_{uu}(g_p-g_m+1)-r_{uu}(g_p-g_m)]+r_{ww}(0)
\end{aligned}$$

(3.23)

Therefore, $E[e_2^{\,2}(n)] = r_{ww}(0)$ for $g_m = g_p$. Now $r_{uu}(0) \geq r_{uu}(n)\forall n$ and for $g_m < g_p$, it may be shown by comparing the sizes of the individual terms in equation (3.23) that $E[e_2^{\,2}(n)] > r_{ww}(0)$ for all values of g_m and g_p (O'Dwyer (1996m)). For $g_m > g_p$, it may also be shown by comparing the sizes of the individual terms that $E[e_2^{\,2}(n)] > r_{ww}(0)$ for all values of g_m and g_p (O'Dwyer (1996m)). Thus, the minimum value of the MSE function occurs at $g_m = g_p$ and the measurement noise has no effect on the estimated process delay index. The only situation that arises for which $E[e_2^{\,2}(n)] = r_{ww}(0)$ for $g_m \neq g_p$ is when the input has a flat autocorrelation function, which corresponds to a constant level input. Thus, any input change is sufficient for correct process delay index estimation, if the process delay index is estimated by determining the minimum of the MSE performance surface.

However, if a gradient method is used to estimate g_p, then an additional restriction that the MSE function must be unimodal about a minimum value when $g_m = g_p$, is imposed. The unimodality of the MSE function in equation (3.23) may be proved by induction; an outline of the inductive proof (provided in full by O'Dwyer (1996m)) is as follows:

(a) For $g_p > g_m$, it is required to prove that the MSE function at $g_m = g_p - 1$ is greater than the MSE function at $g_m = g_p$. It may be proved that this is true, using equation (3.23), provided

$$\left[(1-e^{-T_s/T})^2 + (T_s/T)^2 + T_s(1-e^{-T_s/T})/T\right]\left[r_{uu}(0) - r_{uu}(1)\right] - (T_s/T)(1-e^{-T_s/T})\left[r_{uu}(1) - r_{uu}(2)\right] > 0 \tag{3.24}$$

Simple analysis shows that this expression is always true.

(b) For $g_p > g_m$, it is required to prove that the MSE function at $g_m = g_p - n - 1$ is greater than the MSE function at $g_m = g_p - n$. Applying equation (3.23), it may be proved that this is true, provided

$$\begin{aligned}(1-e^{-T_s/T})^2[r_{uu}(n) - r_{uu}(n+1)] + [\frac{T_s^2}{T^2} + (1-e^{-T_s/T})\frac{T_s}{T}][r_{uu}(0) - r_{uu}(1)]\\ +(1-e^{-T_s/T})\frac{T_s}{T}[nr_{uu}(n) - (2n+1)r_{uu}(n+1) + (n+1)r_{uu}(n+2)] > 0 \end{aligned}\tag{3.25}$$

The condition in equation (3.25) is fulfilled by many excitation signals; one example is a white noise signal.

(c) For $g_p < g_m$, it is required to prove that the MSE function at $g_m = g_p + 1$ is always greater than the MSE function at $g_m = g_p$. Using equation (3.23), it may be proved that this is true, provided

$$2K^2(1-e^{-T_s/T} - T_s/T)^2[r_{uu}(0) - r_{uu}(1)] > 0 \qquad (3.26),$$

which is always true.

(d) For $g_p < g_m$, it is required to prove that the MSE function at $g_m = g_p + n + 1$ is greater than the MSE function at $g_m = g_p + n$. Applying equation (3.23), it may be proved that this is always true, provided

$$(1 - e^{-T_s/T})^2 [r_{uu}(n) - r_{uu}(n+1)] + [\frac{T_s^2}{T^2} - (1 - e^{-T_s/T})\frac{T_s}{T}][r_{uu}(0) - r_{uu}(1)]$$
$$-(1 - e^{-T_s/T})\frac{T_s}{T}[-n r_{uu}(n-1) + (2n+1) r_{uu}(n) - (n+1) r_{uu}(n+1)] > 0 \qquad (3.27)$$

As with equation (3.25), the condition in equation (3.27) is fulfilled by many excitation signals; one example is a white noise signal. □

This theorem is superficially similar to one developed by Elnagger *et al.* (1990a), for application to the estimation of the time delay of a FOLPD process model, in which the time delay is not approximated; these authors do not prove, however, that the corresponding MSE surface is unimodal.

The unimodality of the MSE function (given by equation (3.23)) versus model time delay index is confirmed by representative simulation results given in Figures 3.5 and 3.6. For these simulations, $K_p = K_m = 2.0$, $T_p = T_m = 0.7$ seconds with $g_p = 30$. The normalised MSE (equal to the MSE divided by $r_{uu}(0)$) is plotted versus model time delay index in both cases, with $r_{ww}(0)$ put to zero. The excitation signal used to produce Figure 3.5 is white noise, with the excitation signal used to produce Figure 3.6 being a square wave of period equal to 100 samples.

Considering equations (3.19), (3.20) and (3.21), a block diagram representation of the gradient method to update the model time delay index is shown in Figure 3.7.

Figure 3.5: Normalised MSE vs. time delay index - white noise input

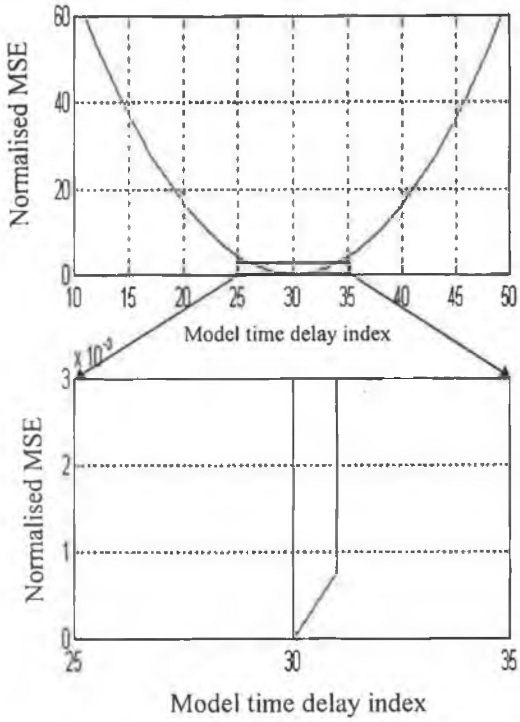


Figure 3.6: Normalised MSE vs. time delay index - square wave input

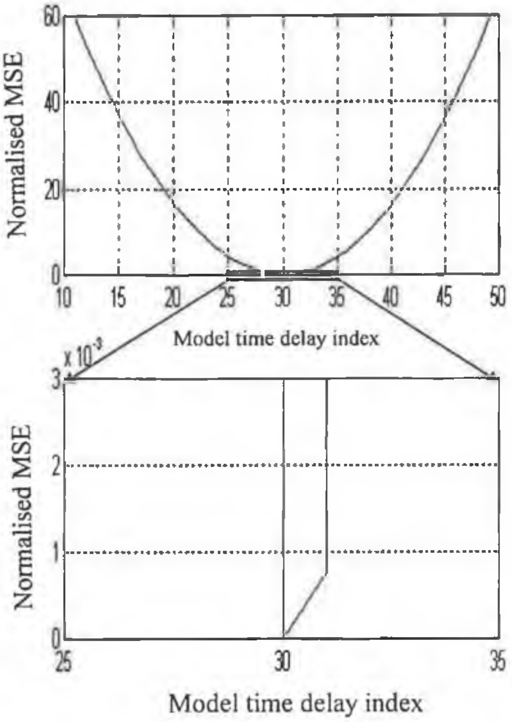
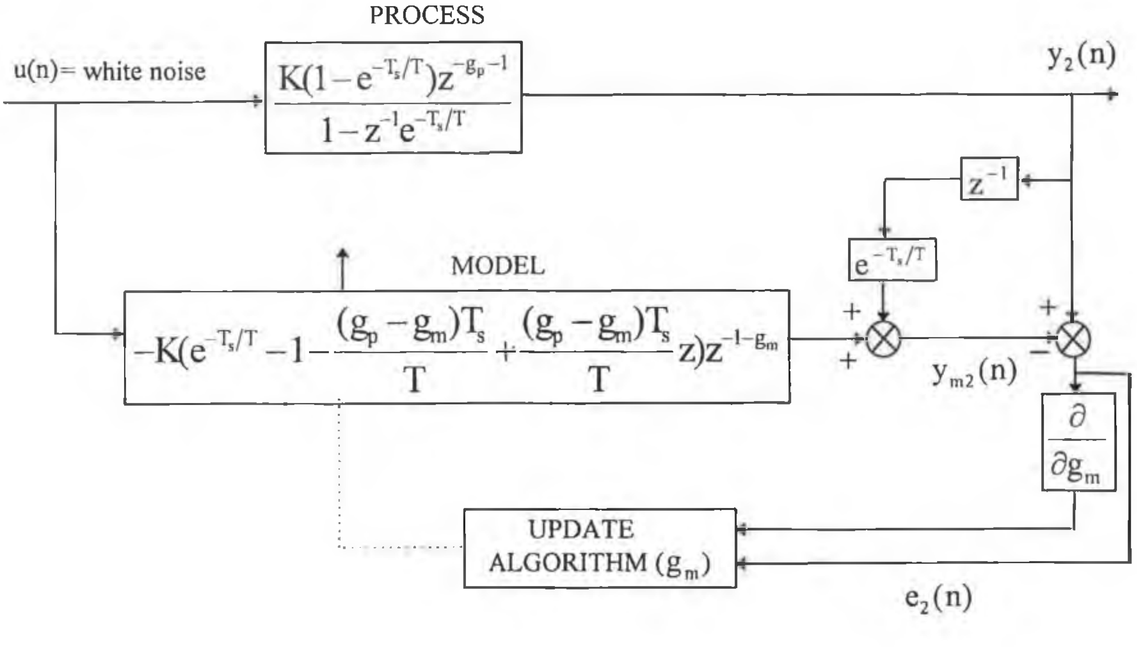


Figure 3.7: Updating of the model time delay index - Case 1.



One representative simulation result corresponding to Theorem 3.2 is given in Figures 3.8a-3.8d. The time delay indices and the process minus model output are plotted against sample number. At the beginning, the starting values of the process and model time delay index were both equalised; a step change was then made to the process time delay index value. In the simulation, the update for the model time delay is a fractional multiple of the sample period; when the addition of these updates exceeds the value of the sample period (in either the positive or negative direction), then an appropriate adjustment is made in the model time delay index, with the update for the model time delay reset to zero. The process and model gain and time constant parameters were put equal to 2.0 and 0.7 seconds, respectively (i.e. the simulation conditions correspond to the conditions taken to calculate the MSE curves in Figures 3.5 and 3.6). The Levenberg-Marquardt gradient algorithm (Ljung (1987)) was used to update the model time delay index; the sample time was defined equal to 0.1 seconds. In the implementation of this algorithm, the starting value of the inverse Hessian matrix was defined equal to $25I$, with $\delta = 0.001$ and $\lambda(n) = 0.95$ (these values were determined from simulation results to be appropriate to the application). Coloured measurement noise, generated by low-pass filtering a white noise signal, was added. The model time delay index was limited in variation to one sample period per iteration (which is a form of filtering on the time delay index value); such filtering was found to be desirable in simulation. Fast convergence to the process time delay index is seen, even in the presence of very substantial coloured measurement noise; this is true if the starting value of the model time delay index is either greater than or less than the process time delay index, as expected from Theorem 3.2. The error, $e_2(n)$, in Figures 3.8b and 3.8d is non-zero due to the presence of the coloured measurement noise.

Figure 3.8a: Time delay index estimate

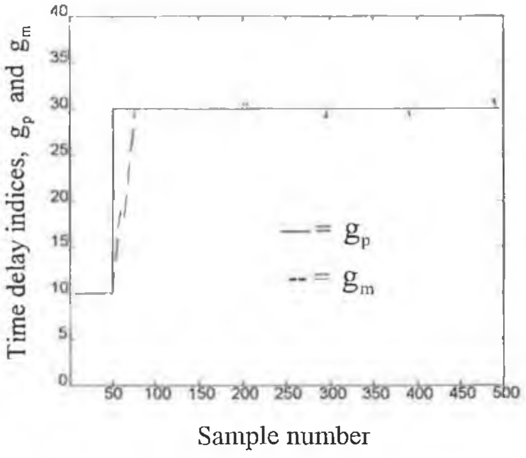


Figure 3.8c: Time delay index estimate

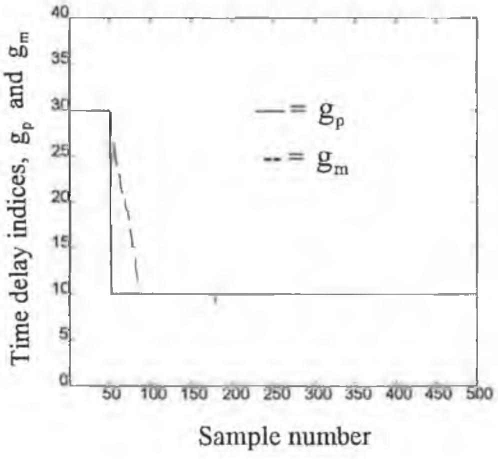


Figure 3.8b: $e_2(n)$ corresponding to Figure 3.8a

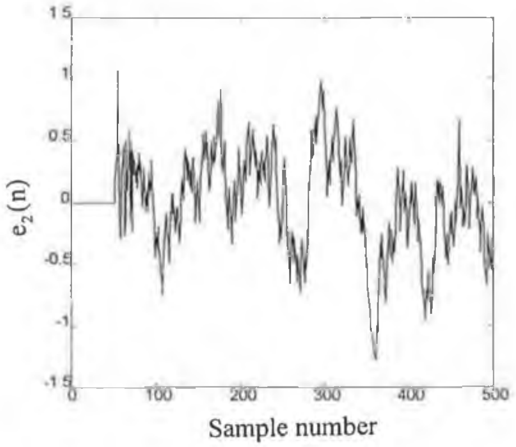
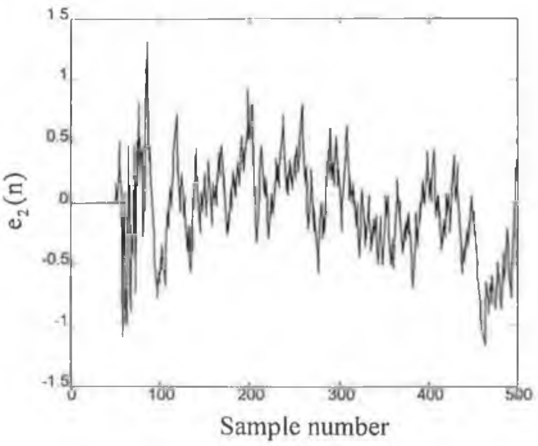


Figure 3.8d: $e_2(n)$ corresponding to Figure 3.8c



The procedures outlined depend on *a priori* knowledge of the process time delay index, g_p (as may be seen clearly in the model in Figure 3.7, for instance). Therefore, the implementation must be regarded as an idealised case. Section 3.4.2 provides the development of a more realistic implementation.

3.4.2 Convergence of the model time delay - Case 2.

3.4.2.1 The time delay as an integer multiple of the sample period

It is necessary to modify the procedure outlined in Section 3.4.1 if *a priori* knowledge of the process time delay index, g_p, is not available (as will normally be the case). One possibility is to calculate the error, $e_3(n)$, based on using a FOLPD process model. The model difference equation in this case is (assuming that the previous process output is used in its calculation, and that $T_p = T_m = T$, $K_p = K_m = K$)

$$y_{m3}(n) = e^{-T_s/T} y_2(n-1) + K(1 - e^{-T_s/T}) u(n - g_m - 1) \qquad (3.28)$$

with $y_2(n)$ given by equation (3.19). Therefore, from equations (3.19) and (3.28), $e_3(n) = y_2(n) - y_{m3}(n)$ is given by

$$e_3(n) = K(1 - e^{-T_s/T})[u(n - g_p - 1) - u(n - g_m - 1)] + w(n) \tag{3.29}$$

The partial derivative of the error with respect to the time delay variation may then be calculated by using a first order Taylor's series approximation for the time delay variation. This error, $e_2(n)$, is given by equation (3.21); the corresponding partial derivative mentioned above is

$$\frac{\partial e_2(n)}{\partial(g_p - g_m)} = \frac{KT_s}{T}\left[u(n - g_m) - u(n - g_m - 1)\right] \tag{3.30}$$

The update vector (for updating the model time delay - equation (2.4)), which depends on the product of the error $(e_3(n))$ multiplied by the partial derivative of the error with respect to the time delay variation $(\partial e_2(n)/\partial(g_p - g_m))$ is then independent of g_p. The cost function that approximately corresponds to this update vector will be referred to as the mean of the product of the errors (MPE) function; this function is defined as $E[e_2(n)e_3(n)]$ in this case. The update vector that exactly corresponds to this cost function depends on $e_3(n)[\partial e_2(n)/\partial(g_p - g_m)]$ and $e_2(n)[\partial e_3(n)/\partial(g_p - g_m)]$. It is assumed that $e_3(n)[\partial e_2(n)/\partial(g_p - g_m)] \approx e_2(n)[\partial e_3(n)/\partial(g_p - g_m)]$. This is a reasonable assumption, bearing in mind that the time delay variation, which is approximated by a first order Taylor's series approximation, is assumed to be small.

The MPE cost function will equal the MSE cost function at $g_p = g_m$ (when equation (3.20) reduces to equation (3.28)).

If any other approximation to the time delay variation is used rather than a first order Taylor's series approximation, then it may be shown (O'Dwyer (1996e)) that the partial derivative of the error with respect to the time delay variation is a function of g_p. Thus, if g_p is unknown *a priori*, then a first order Taylor's series approximation for the time delay variation is the only approximation of interest.

It is desired to prove convergence of the model time delay index to the process time delay index, with the process time delay index being unknown, but with the other model parameters being known *a priori.*

Theorem 3.3: For a first order discrete stable system of known gain and time constant, then the MPE performance surface versus model time delay index is unimodal, with a minimum value of the MPE occurring when the model time delay index equals the process time delay index, under the following conditions:

(a) The time delay variation is approximated by a first order Taylor's series approximation

(b) The measurement noise is uncorrelated with the process input

(c) The resolution on the process time delay is assumed to be equal to one sample period

(d) The error is calculated based on using a FOLPD process model; the partial derivative of the error with respect to the time delay variation is calculated based on using the first order Taylor's series approximation for the time delay variation and

(e) The process time delay index is greater than the model time delay index, as the model time delay index converges.

Proof: The process difference equation, $y_2(n)$, is given by equation (3.19). The model difference equation, $y_{m3}(n)$, is given by equation (3.28). The model difference equation for calculating the partial derivative of the error with respect to the time delay variation, $y_{m2}(n)$, is given by equation (3.20). The expressions $e_2(n) = y_2(n) - y_{m2}(n)$ and $e_3(n) = y_2(n) - y_{m3}(n)$ are given by equations (3.21) and (3.29), respectively. The MPE performance surface, $E[e_2(n)e_3(n)]$, may then be calculated, using a procedure similar to that used in equations (3.22) and (3.23), to be equal to (O'Dwyer (1996m))

$$K^2(1 - e^{-T_s/T})\frac{(g_p - g_m)T_s}{T}[r_{uu}(0) - r_{uu}(1) + r_{uu}(g_p - g_m + 1) - r_{uu}(g_p - g_m)] + r_{ww}(0)$$
$$+2K^2(1 - e^{-T_s/T})^2[r_{uu}(0) - r_{uu}(g_p - g_m)] \qquad (3.31)$$

Therefore, $E[e_2(n)e_3(n)] = r_{ww}(0)$ for $g_m = g_p$. It may be shown by comparing the sizes of the individual terms in equation (3.31) that $E[e_2(n)e_3(n)] > r_{ww}(0)$ for $g_p > g_m$ only (O'Dwyer (1996m)).

Thus, the minimum value of $E[e_2(n)e_3(n)]$ occurs at $g_m = g_p$ (when g_m is restricted to be less than or equal to g_p) and the measurement noise has no effect on the estimated process delay value. If $g_p > g_m$, then, from equation (3.31), the only situation that arises for which $E[e_2(n)e_3(n)] = r_{ww}(0)$ for $g_m \neq g_p$ is when the input has a flat autocorrelation function, which corresponds to a constant level input. Thus, any input change is sufficient for correct process delay index estimation, provided that the required condition on g_m is fulfilled, if the process delay index is estimated by determining the minimum of the MPE performance surface.

However, if a gradient method is used to estimate g_p, then an additional restriction that the MPE function must be unimodal for $g_p > g_m$, with a minimum MPE value occurring at $g_m = g_p$, is imposed. The unimodality of the MPE function for $g_p > g_m$ may be proved by induction; an outline of the inductive proof (provided in full by O'Dwyer (1996m)) is as follows:

It may be proved that the MPE function at $g_m = g_p - 1$ is greater than the MPE function at $g_m = g_p$ (using equation (3.31)), provided that

$$2(1 - e^{-T_s/T})[r_{uu}(0) - r_{uu}(1)] + \frac{T_s}{T}[r_{uu}(0) - 2r_{uu}(1) + r_{uu}(2)] > 0 \qquad (3.32)$$

Applying equation (3.31), it may be proved that the MPE function at $g_m = g_p - n - 1$ is greater than the MPE function at $g_m = g_p - n$, provided that

$$2(1 - e^{-T_s/T})[r_{uu}(n) - r_{uu}(n+1)] + \frac{T_s}{T}[r_{uu}(0) - r_{uu}(1)]$$
$$+ \frac{T_s}{T}[nr_{uu}(n) - (2n+1)r_{uu}(n+1) + (n+1)r_{uu}(n+2)] > 0 \qquad (3.33)$$

Both of the conditions in equations (3.32) and (3.33) are fulfilled by many excitation signals; one example is a white noise signal. □

The behaviour of the MPE function (given by equation (3.31)) versus model time delay index is confirmed by Figures 3.9 and 3.10, in representative simulation results. For these simulations, $K_p = K_m = 2.0$, $T_p = T_m = 0.7$ seconds and $g_p = 30$; these conditions are identical to those used to calculate the normalised MSE curves in Figures 3.5 and 3.6. The normalised MPE (equal to the MPE divided by $r_{uu}(0)$) is plotted versus model time delay index in both cases, with $r_{ww}(0)$ put to zero. The excitation signal used to produce Figure 3.9 is white noise, with the excitation signal used to produce Figure 3.10 being a square wave of period equal to 100 samples. These plots show that the MPE performance surface is greater than $r_{ww}(0)$ for $g_p > g_m$ only, and that when the conditions in equations (3.32) and (3.33) are fulfilled, the MPE function is unimodal for $g_p > g_m$, with a minimum MPE value occurring at $g_m = g_p$.

Figure 3.9: Normalised MPE vs. time delay index - white noise input

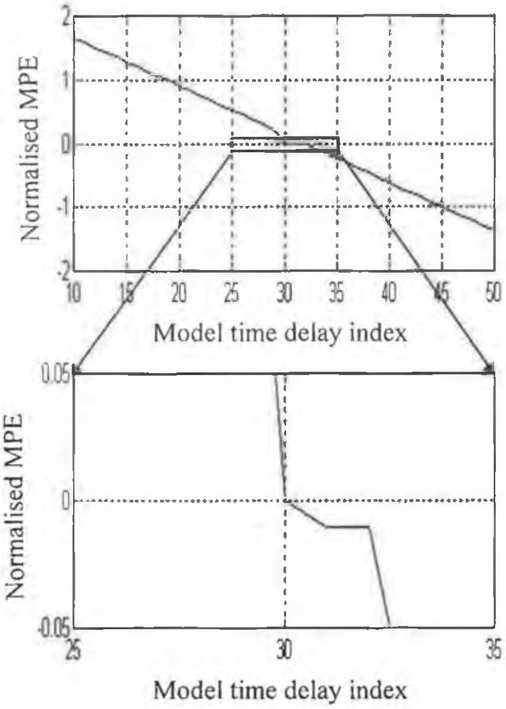
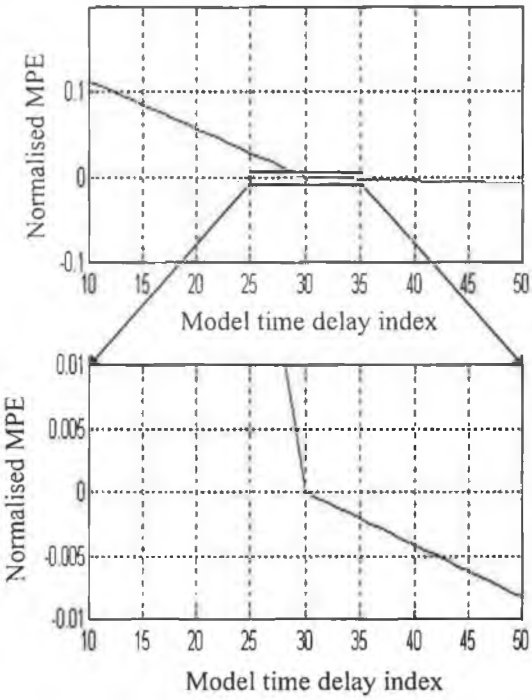


Figure 3.10: Normalised MPE vs. time delay index - square wave input



Considering equations (3.19), (3.20), (3.21), (3.28) and (3.29), a block diagram representation of the gradient method to update the model time delay index is shown in Figure 3.11.

Representative simulation results corresponding to Theorem 3.3 are given in Figures 3.12 and 3.13, with the time delay indices and the process minus model output plotted against sample number. The simulation conditions are identical to those used in Section 3.4.1 (and thus correspond to the conditions taken to calculate the MPE curves in Figures 3.9 and 3.10), with the addition that in the simulation in which a square wave is the excitation signal, the learning rate, μ, for the time delay is put to 10 and filtering on the time delay update is employed; these conditions were determined to be appropriate for the application.

Figure 3.11: Updating of the model time delay index - Case 2

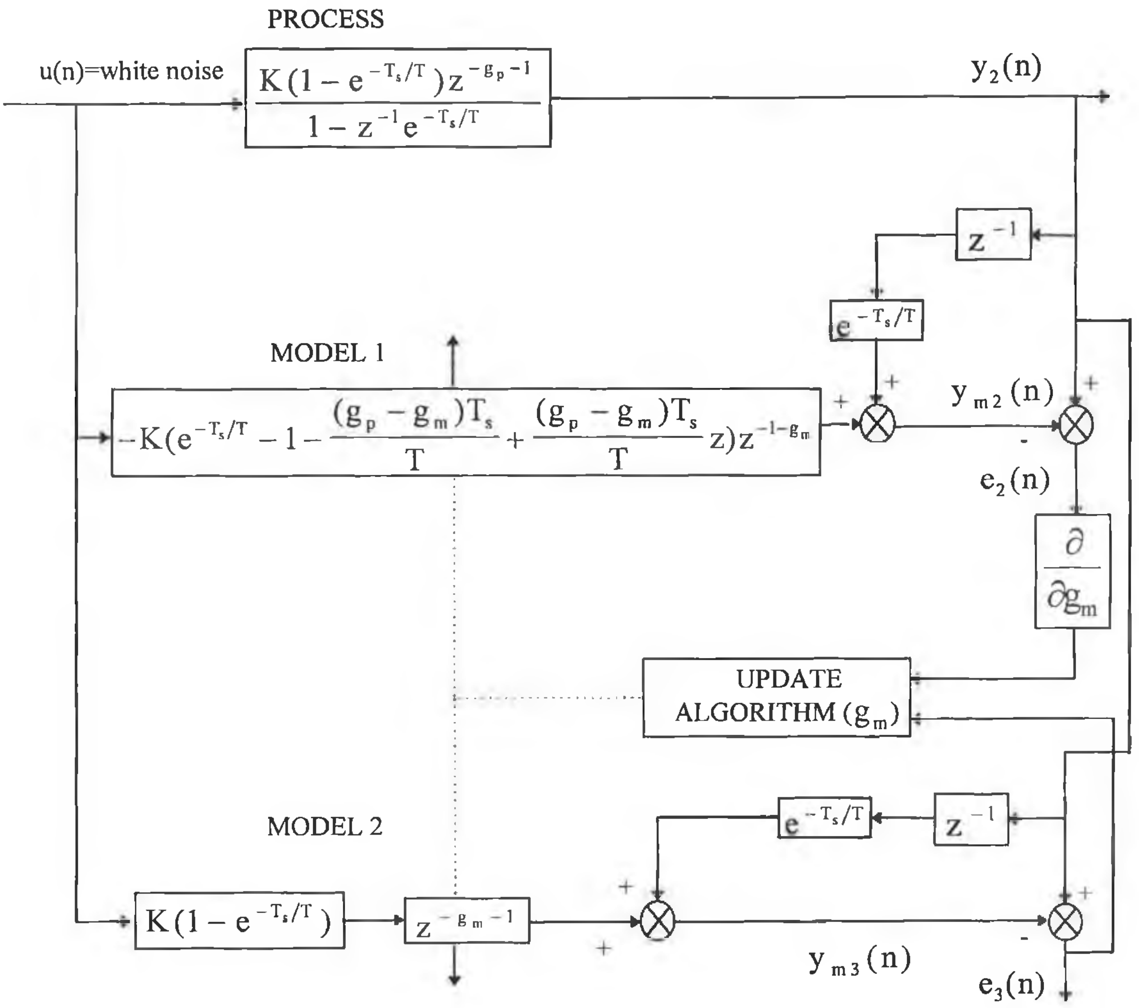
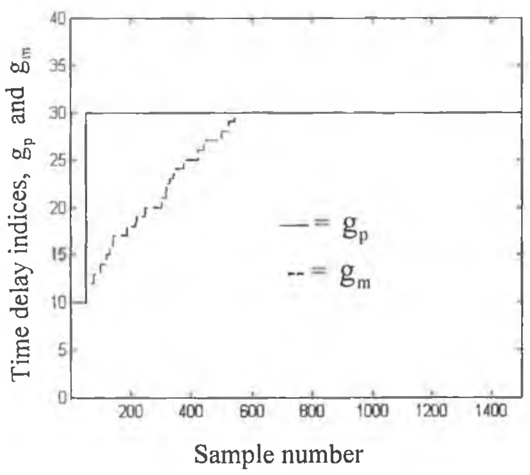


Figure 3.12a: Time delay index estimate-white noise excitation

Figure 3.13a: Time delay index estimate-square wave excitation



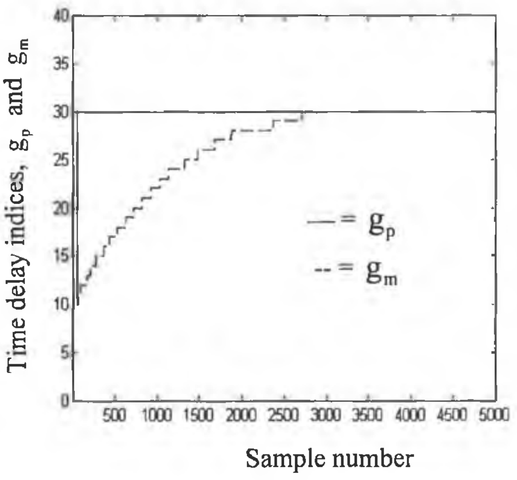


Figure 3.12b: $e_3(n)$ corresponding to Figure 3.12a

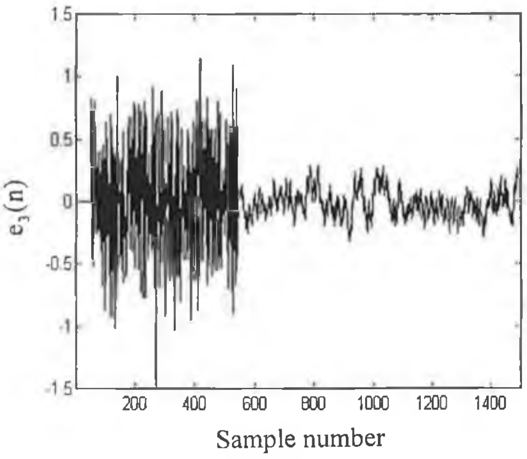
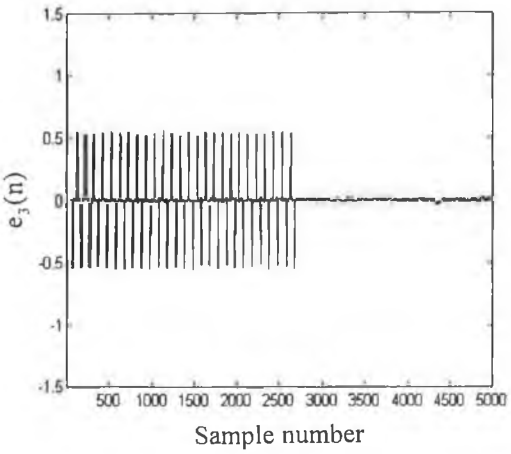


Figure 3.13b: $e_3(n)$ corresponding to Figure 3.13a



Good convergence to the process time delay index is seen for $g_p > g_m$. Other supplementary simulation results show no convergence to the process time delay index when $g_p < g_m$. This verifies Theorem 3.3. However, the nature of the MPE functions (Figures 3.9 and 3.10) mean that convergence of the model time delay index could not be guaranteed, as the MPE goes negative when $g_m > g_p$. Convergence could only be guaranteed if g_m is always less than or equal to g_p. The convergence of the model time delay index in the simulations taken is due to the manner in which the parameter is being updated in the simulation, which tends to prevent g_m going greater than g_p; as Figures 3.12b and 3.13b show, estimation is possible in the presence of coloured measurement noise only when such noise is at a low amplitude.

3.4.2.2 The time delay as a real multiple of the sample period

Theorem 3.3 in the previous section dealt with the estimation of time delays that are integer multiples of the sample period. For the estimation of time delays that are real multiples of the sample period (and assuming $T_p = T_m = T$, $K_p = K_m = K$), the FOLPD process difference equation is given as (O' Dwyer (1996e)):

$$y_3(n) = e^{-T_s/T} y_3(n-1) + K(1 - e^{g_b T_s/T}) u(n - g_p) + K(e^{g_b T_s/T} - e^{-T_s/T}) u(n - g_p - 1) + w(n) \qquad (3.34)$$

with g_b = process time delay minus the process time delay index.

The corresponding model difference equation (assuming the previous process output is used in its calculation) is

$$y_{m4}(n) = e^{-T_s/T}y_3(n-1) + K(1 - e^{g_a T_s/T})u(n-g_m) + K(e^{g_a T_s/T} - e^{-T_s/T})u(n-g_m-1) \tag{3.35}$$

with g_a = model time delay minus model time delay index.

The model difference equation for calculating the partial derivative of the error with respect to the time delay variation (and assuming that the previous process output is used in its calculation) is

$$\begin{aligned} y_{m5}(n) = e^{-T_s/T}y_3(n-1) - \frac{K(g_p - g_m + g_b - g_a)T_s}{T}u(n-g_m) \\ -K[e^{-T_s/T} - 1 - \frac{(g_p - g_m + g_b - g_a)T_s}{T}]u(n-g_m-1) \end{aligned} \tag{3.36}$$

This equation may be deduced from equation (3.20). Therefore, from equations (3.34) and (3.35),

$$\begin{aligned} e_4(n) = y_3(n) - y_{m4}(n) = K(1 - e^{g_b T_s/T})u(n-g_p) + K(e^{g_b T_s/T} - e^{-T_s/T})u(n-g_p-1) \\ -K(1 - e^{g_a T_s/T})u(n-g_m) - K(e^{g_a T_s/T} - e^{-T_s/T})u(n-g_m-1) + w(n) \end{aligned} \tag{3.37}$$

and, from equations (3.34) and (3.36),

$$\begin{aligned} e_5(n) = y_3(n) - y_{m5}(n) = K(1 - e^{g_b T_s/T})u(n-g_p) + K(e^{g_b T_s/T} - e^{-T_s/T})u(n-g_p-1) \\ + \frac{KT_s(g_p - g_m + g_b - g_a)}{T}u(n-g_m) + K[e^{-T_s/T} - 1 - \frac{T_s(g_p - g_m + g_b - g_a)}{T}]u(n-g_m-1) + w(n) \end{aligned} \tag{3.38}$$

The MPE performance surface, $E[e_4(n)e_5(n)]$, may then be calculated, using a procedure similar to that used in equations (3.22) and (3.23), to be equal to (O'Dwyer (1996m))

$$
\begin{aligned}
&K^2[(1-e^{g_b T_s/T})^2+(e^{g_b T_s/T}-e^{-T_s/T})^2+(1-e^{-T_s/T})(e^{g_a T_s/T}-e^{-T_s/T})]r_{uu}(0)\\
&\quad -K^2(1+e^{-T_s/T}-2e^{g_a T_s/T})\frac{T_s}{T}(g_p-g_m+g_b-g_a)r_{uu}(0)+\\
&\quad K^2[2(1-e^{g_b T_s/T})(e^{g_b T_s/T}-e^{-T_s/T})+(1-e^{-T_s/T})(1-e^{g_a T_s/T})\\
&\quad -(2e^{g_a T_s/T}-1-e^{-T_s/T})\frac{T_s}{T}(g_p-g_m+g_b-g_a)]r_{uu}(1)\ +\\
&\quad K^2[\frac{(g_p-g_m+g_b-g_a)T_s}{T}(1-2e^{g_b T_s/T}+e^{-T_s/T})-(1-e^{g_b T_s/T})(1-e^{g_a T_s/T})\\
&\qquad -\ (e^{g_b T_s/T}-e^{-T_s/T})(1-2e^{-T_s/T}+e^{g_a T_s/T})]r_{uu}(g_p-g_m)\\
&-K^2(1-e^{g_b T_s/T})[\frac{(g_p-g_m+g_b-g_a)T_s}{T}+(1-2e^{-T_s/T}+e^{g_a T_s/T})]r_{uu}(g_p-g_m-1)\ +\\
&\quad K^2(e^{g_b T_s/T}-e^{-T_s/T})[\frac{(g_p-g_m+g_b-g_a)T_s}{T}-(1-e^{g_a T_s/T})]r_{uu}(g_p-g_m+1)+r_{ww}(0)
\end{aligned}
\tag{3.39}
$$

Now, using equation (3.39), it may be shown that $E[e_4(n)e_5(n)] = r_{ww}(0)$ if $g_p = g_m$ and $g_b = g_a$. The behaviour of the MPE function versus model time delay is given by Figures 3.14, 3.15, 3.16 and 3.17, in representative simulation results. For these simulations, $K_p = K_m = 2.0$, $T_p = T_m = 0.7$ seconds and $g_p = 5$, with the time delay taken in intervals of 0.01 times the sample period. The normalised MPE (equal to the MPE divided by $r_{uu}(0)$) is plotted versus model time delay index for $g_b = 0.0$ and $g_b = 0.5$ in Figures 3.14 and 3.15, when the excitation signal to the process is white noise. The normalised MPE is plotted versus model time delay index for $g_b = 0.0$ and $g_b = 0.5$ in Figures 3.16 and 3.17, when the excitation signal to this process is a square wave of period equal to 100 samples. In Figures 3.14 and 3.16, the normalised MPE calculated from equation (3.31) is superimposed on the plots for comparison purposes. For all simulations, $r_{ww}(0)$ is put to zero.

Figures 3.14-3.17 show the true multimodal nature of the MPE function versus model time delay when the time delay is a real multiple of the sample period. The estimation of the real value of the process time delay is impossible using gradient methods.

Figure 3.14: Normalised MPE vs. time delay index - white noise excitation - $g_b = 0.0$

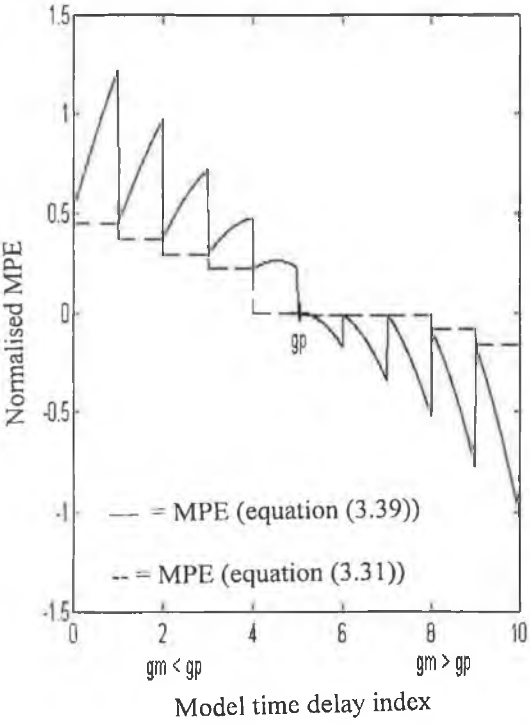


Figure 3.15: Normalised MPE vs. time delay index - white noise excitation - $g_b = 0.5$

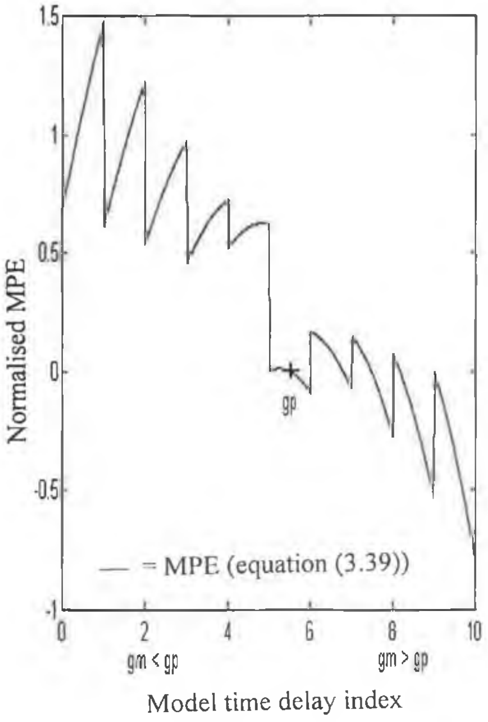


Figure 3.16: Normalised MPE vs. time delay index - square wave excitation - $g_b = 0.0$

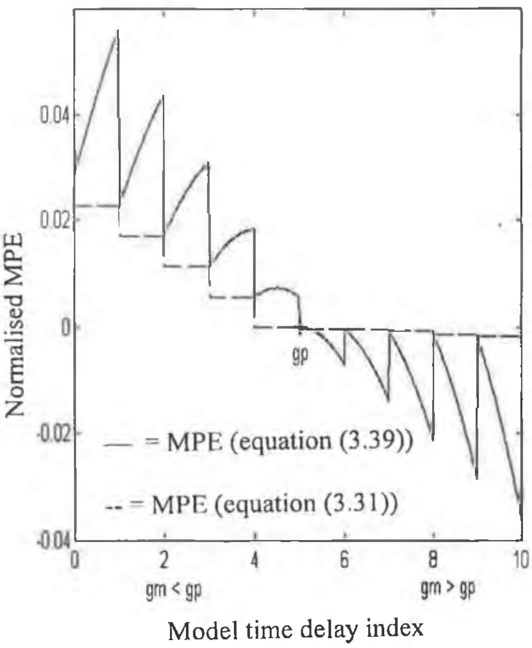
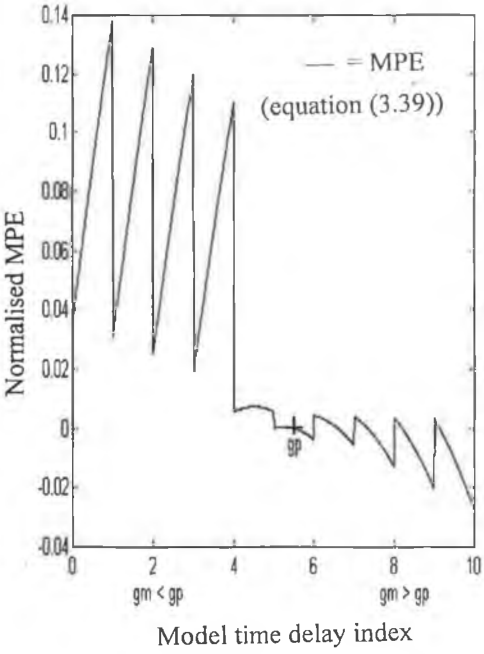


Figure 3.17: Normalised MPE vs. time delay index - square wave excitation - $g_b = 0.5$



3.4.2.3 Conclusions

In summary, the gradient method will allow the estimation of process time delays that are integer multiple of the sample period, in the case where the process time delay is the only unknown parameter, provided the process time delay is always greater than the model time delay, as the model time delay converges to the process time delay. An alternative non-gradient method that involves estimating the process time delay index by determining the minimum positive value of the MPE performance surface would allow the estimation of the process time delay index under the same conditions as the gradient method. Unfortunately, it is not possible to estimate process time delays that are not integer multiples of the sample period using the gradient method, though it appears from Figures 3.14-3.17 that it may be possible to do so using a non-gradient method based on determining the minimum positive value of the MPE performance surface (at least for $g_p > g_m$).

3.4.3 Convergence of the model time delay - Case 3

In Theorem 3.3, the previous process output is used to calculate the new model output. It was decided to investigate the convergence pattern of the model time delay to the process time delay if the previous model output is used to calculate the new model output. The time delay is not assumed known *a priori*.

Theorem 3.4: For a first order discrete stable system of known gain and time constant, the MPE performance surface versus model time delay index is unimodal, with a minimum value of the MPE occurring when the model time delay index equals the process time delay index, under the following conditions:

(a) The time delay variation is approximated by a first order Taylor's series approximation

(b) The measurement noise is uncorrelated with the process input and output

(c) The resolution on the process time delay is assumed to be equal to one sample period

(d) The error is calculated based on using a FOLPD process model; the partial derivative of the error with respect to the time delay variation is calculated based on using the first order Taylor's series approximation for the time delay variation

(e) The conditions provided in the theorem are observed on the model parameters and

(f) The previous model output is used to calculate the new model output.

Proof: The process difference equation, $y_2(n)$, is given by equation (3.19). The corresponding FOLPD model difference equation is (assuming the previous model output is used in its calculation)

$$y_{m6}(n) = e^{-T_s/T} y_{m6}(n-1) + K(1 - e^{-T_s/T})u(n - g_m - 1) \tag{3.40}$$

The model difference equation for calculating the partial derivative of the error with respect to the time delay variation (and assuming that the previous model output is used in its calculation) is

$$y_{m7}(n) = e^{-T_s/T} y_{m7}(n-1) - \frac{K(g_p - g_m)T_s}{T} u(n - g_m) - K(e^{-T_s/T} - 1 - \frac{(g_p - g_m)T_s}{T})u(n - g_m - 1) \tag{3.41}$$

Therefore, from equations (3.19) and (3.40),

$$\begin{aligned} e_6(n) = y_2(n) - y_{m6}(n) = e^{-T_s/T}[y_2(n-1) - y_{m6}(n-1)] \\ + K(1 - e^{-T_s/T})[u(n - g_p - 1) - u(n - g_m - 1)] + w(n) \end{aligned} \tag{3.42}$$

and, from equations (3.19) and (3.41),

$$\begin{aligned} e_7(n) = y_2(n) - y_{m7}(n) = e^{-T_s/T}[y_2(n-1) - y_{m7}(n-1)] + K(1 - e^{-T_s/T})u(n - g_p - 1) \\ + K[\frac{(g_p - g_m)T_s}{T} u(n - g_m) + (e^{-T_s/T} - 1 - \frac{(g_p - g_m)T_s}{T})u(n - g_m - 1)] + w(n) \end{aligned} \tag{3.43}$$

The MPE performance surface, $E[e_6(n)e_7(n)]$, (which will approximately correspond to the update vector formed from the product of $e_6(n)$ and $\partial e_7(n)/\partial(g_p - g_m)$, and will thus be independent of g_p) may then be calculated, using a procedure similar to that used in equations (3.22) and (3.23), to be equal to (O'Dwyer (1996m)

$$K^2(1-e^{-T_s/T})\frac{(g_p-g_m)T_s}{T}[r_{uu}(0)-r_{uu}(1)+r_{uu}(g_p-g_m+1)-r_{uu}(g_p-g_m)]+r_{ww}(0)$$
$$+2K^2(1-e^{-T_s/T})^2[r_{uu}(0)-r_{uu}(g_p-g_m)]+e^{-2T_s/T}[r_{y_2y_2}(0)+r_{y_{m6}y_{m7}}(0)-r_{yy_{m6}}(0)-r_{yy_{m7}}(0)]$$
$$-Ke^{-T_s/T}(1-e^{-T_s/T})r_{uy_{m6}}(g_p)+Ke^{-T_s/T}(1-e^{-T_s/T})[r_{uy_{m7}}(g_m)-r_{uy_{m7}}(g_p)]$$
$$-e^{-T_s/T}\frac{KT_s}{T}(g_p-g_m)r_{uy_{m6}}(g_m-1)-Ke^{-T_s/T}[e^{-T_s/T}-1-\frac{(g_p-g_m)T_s}{T}]r_{uy_{m6}}(g_m)$$
$$+2e^{-T_s/T}K(1-e^{-T_s/T})[r_{uy_2}(g_p)-r_{uy_2}(g_m)]+e^{-T_s/T}K(g_p-g_m)\frac{T_s}{T}[r_{uy_2}(g_m-1)-r_{uy_2}(g_m)]$$
(3.44)

with $r_{uy_x}(n)$ being the cross-correlation function between u(n) and $y_x(n)$, $r_{y_2y_2}(n)$ being the autocorrelation function of $y_2(n)$ and $r_{y_{x1}y_{x2}}(n)$ being the cross-correlation function between $y_{x1}(n)$ and $y_{x2}(n)$. These terms may be calculated as follows (O'Dwyer (1996m)):

$$r_{y_2y_2}(0)=\frac{K^2(1-e^{-T_s/T})^2}{1-e^{-2T_s/T}}r_{uu}(0) \quad (3.45)$$

$$r_{yy_{m6}}(0)=\frac{Ke^{-T_s/T}(1-e^{-T_s/T})r_{uy_{m6}}(g_p)+K^2(1-e^{-T_s/T})^2r_{uu}(g_p-g_m)}{(1-e^{-2T_s/T})} \quad (3.46)$$

$$r_{yy_{m7}}(0)=\frac{Ke^{-T_s/T}(1-e^{-T_s/T})r_{uy_{m7}}(g_p)-K^2[e^{-T_s/T}-1-\frac{(g_p-g_m)T_s}{T}](1-e^{-T_s/T})r_{uu}(g_p-g_m)}{(1-e^{-2T_s/T})}$$
(3.47)

$$r_{y_{m6}y_{m7}}(0)=\frac{Ke^{-T_s/T}(1-e^{-T_s/T})[r_{uy_{m7}}(g_m)+r_{uy_{m6}}(g_m)]-K^2[e^{-T_s/T}-1-\frac{(g_p-g_m)T_s}{T}](1-e^{-T_s/T})r_{uu}(0)}{(1-e^{-2T_s/T})}$$
$$+\frac{K(g_p-g_m)T_s}{T}e^{-T_s/T}[r_{uy_{m6}}(g_m)-r_{uy_{m6}}(g_m-1)]/(1-e^{-2T_s/T}) \quad (3.48)$$

and $r_{uy_{m6}}(n)=r_{uy_{m6}}(n-1)+K(e^{-T_s/T})^{n-1}(1-e^{-T_s/T})r_{uu}(0)$ (3.49)

$$r_{uy_{m7}}(n) = e^{-T_s/T} r_{uy_{m7}}(n-1) - K[e^{-T_s/T} - 1 - \frac{(g_p - g_m)T_s}{T}] r_{uu}(0) \tag{3.50}$$

$$r_{uy_2}(n) = e^{-(n-1)T_s/T} K_p (1 - e^{-T_s/T_p}) r_{uu}(0)\text{, } n > g_p \text{ and } r_{uy_2}(n) = 0 \text{ otherwise.} \tag{3.51}$$

If a gradient method is used to determine g_p, when $g_p > g_m$, then the MPE function (equation (3.44)) must be unimodal for $g_p > g_m$, with a minimum MPE value occurring at $g_m = g_p$. The unimodality of the MPE for $g_p > g_m$ may be proved by induction; an outline of the inductive proof (provided in full by O'Dwyer (1996m)) is as follows:

It may be proved, using equation (3.44), that the MPE function at $g_m = g_p - 1$ is always greater than the MPE function at $g_m = g_p$. Similarly, it may be proved that the MPE function at $g_m = g_p - n - 1$ is greater than the MPE function at $g_m = g_p - n$, provided that

$$1 - e^{-T_s/T} > e^{-(g_p - n - 1)T_s/T}[1 - (g_p - 2)e^{-T_s/T}(1 - e^{-T_s/T})] \tag{3.52}$$

This is a sufficient condition. □

The behaviour of the MPE function, given by equation (3.44), versus time delay index is confirmed by Figures 3.18 and 3.19, in representative simulation results. For these simulations, $K_p = K_m = 2.0$, $T_p = T_m = 0.7$ seconds and $g_p = 30$. The normalised MPE (equal to the MPE divided by $r_{uu}(0)$) is plotted versus model time delay index in both cases, with $r_{ww}(0)$ put to zero. The excitation signal in both cases is white noise, with the sample period taken to equal 0.1 seconds for Figure 3.18 and 0.02 seconds for Figure 3.19 (this means that the MPE function in Figure 3.18 may be directly compared with that in Figure 3.9). Non-unimodal behaviour is seen in the latter case (for $g_p > g_m$) at values of n when the conditions for convergence are violated.

It is obvious from Figures 3.18 and 3.19 (without the necessity of a proof by induction) that convergence of the model time delay index to the process time delay index is not possible when $g_p < g_m$.

Figure 3.18: Normalised MPE vs. time delay index - white noise excitation - $T_s = 0.1$ seconds

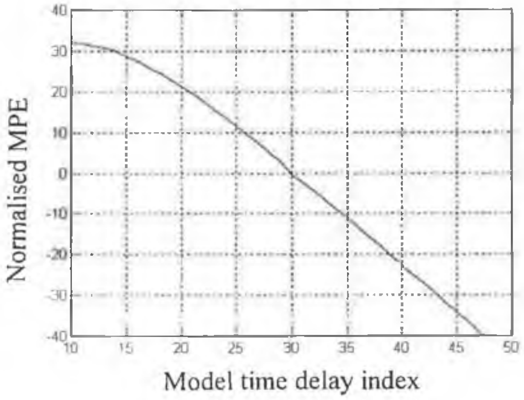
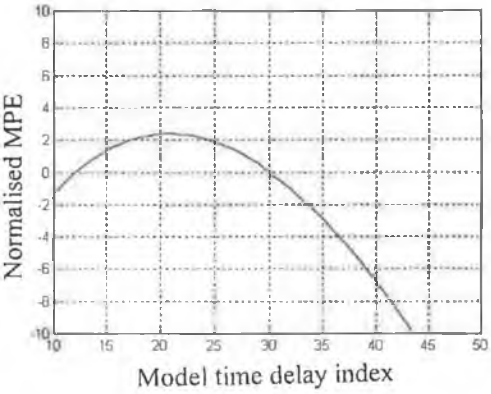


Figure 3.19: Normalised MPE vs. time delay index - white noise excitation - $T_s = 0.02$ seconds



Considering equations (3.19), (3.40), (3.41), (3.42) and (3.43), the block diagram representation of the gradient method to update the model time delay index is as shown in Figure 3.20.

A representative simulation result corresponding to Theorem 3.4 is shown in Figures 3.21a and 3.21b, with the time delay indices and the process minus the model output plotted against sample number. The simulation conditions are identical to those used in Section 3.4.1, and thus correspond to the conditions taken to calculate the MPE curve in Figure 3.18; the results may be directly compared to those shown in Figures 3.12a and 3.12b. The results in Figures 3.21a and 3.21b show that convergence of the model time delay index to the process time delay index is possible if the conditions for convergence are fulfilled. However, the nature of the MPE function (Figure 3.18) means that such convergence could not be guaranteed, and the convergence of the model time delay index in this case is due to the manner in which the parameter is being updated in the simulation (which prevents g_m going greater than g_p). Figure 3.21b shows the low level of coloured measurement noise used in the simulation.

Figure 3.20: Updating of the model time delay index - Case 3.

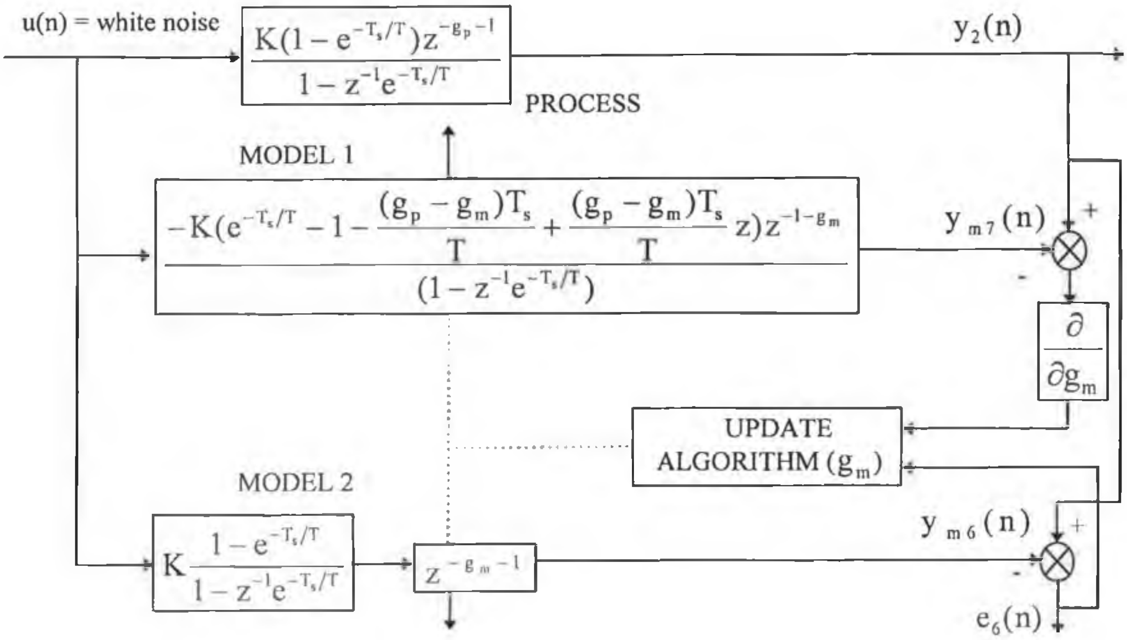
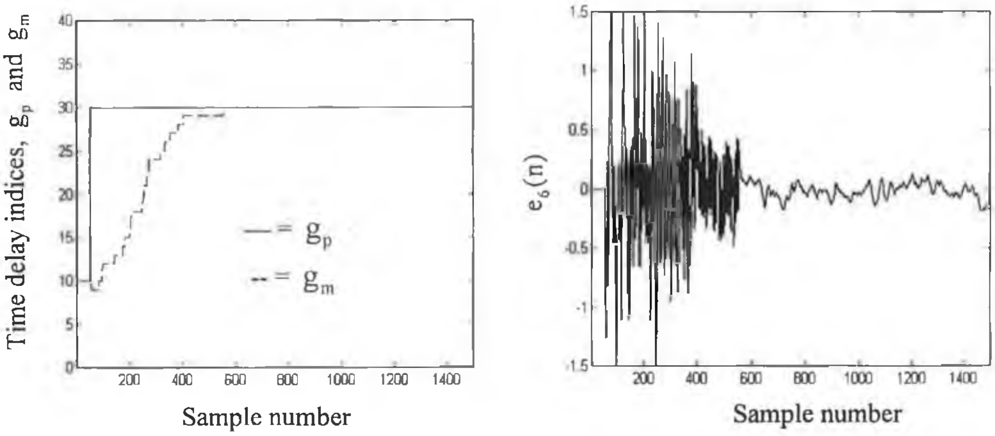


Figure 3.21a: Time Delay Index estimate - white noise excitation

Figure 3.21b: $e_6(n)$ corresponding to Figure 3.21a



Overall, the use of the previous model output to calculate the new model output, rather than the use of the previous process output to calculate the new model output, does not appear to be beneficial, because of the narrower conditions for convergence of the latter implementation, compared to the former implementation.

3.5 Convergence of the full parameter set

This section of the chapter will consider the use of the gradient algorithm for updating all of the model parameter values. The gradient algorithm used is a function of the error between the process and the output, and the partial derivative of the error between the process and the output. The following cases are considered, to comprehensively cover the implementations possible:

1. The error and the partial derivative of the error with respect to the time delay variation are calculated by using a first order Taylor's series approximation for the time delay variation. This will be referred to as Case 1 in subsequent work in this section of the chapter, and corresponds to Case 1 in Section 3.4, where only the model time delay index is updated. As in Section 3.4, this is an idealised implementation, as the process time delay is assumed known *a priori* (and is assumed to be an integer multiple of the sample period).

2. The partial derivative of the error with respect to the time delay variation is calculated by using a first order Taylor's series approximation for the time delay variation, and the error is calculated based on using a FOLPD process model. This will be referred to as Case 2 in subsequent work in this section of the chapter (and it corresponds to Case 2 in Section 3.4, so that it provides a more realistic implementation than Case 1 above, as the process time delay is not assumed known *a priori*). In this case, the following conditions are considered:

(a) The process time delay is an integer multiple of the sample period - white noise input.

(b) The process time delay is a real multiple of the sample period - white noise input.

(c) The process time delay is an integer multiple of the sample period - square wave input.

(d) The process time delay is a real multiple of the sample period - square wave input.

In all cases, the model gain and time constant are updated assuming a FOLPD process model.

The case where the previous model output is used to calculate the new model output is not considered in detail, as the results in Section 3.4.3 revealed that there was no benefit, when the model time delay index was being updated, in implementing such a procedure compared to the procedures in Case 2, Section 3.4.2 (when the previous process output is used to calculate the new model output).

3.5.1 Convergence of the full parameter set - Case 1

Theorem 3.5: For a first order discrete stable system of unknown parameters, the MSE performance surface versus model time delay index is unimodal, with a minimum value of the MSE occurring when the model time delay index equals the process time delay index, under the following conditions:

(a) The time delay variation is approximated by a first order Taylor's series approximation

(b) The measurement noise is uncorrelated with the process input and output

(c) The resolution on the process time delay is assumed to be equal to one sample period

(d) The error and the partial derivative of the error with respect to the time delay variation are calculated based on using the first order Taylor's series approximation for the time delay variation

(e) The excitation signal to the process and model is assumed to be white noise

(f) The model gain and time constant are updated based on using a FOLPD process model and

(g) The conditions provided in the theorem are observed on the model parameters.

Proof: The process difference equation, $y_1(n)$, is given by equation (3.12). The corresponding model difference equation, $y_{m1}(n)$, is given by equation (3.13) (assuming the previous process output has been used in its calculation). The model difference equation, calculated by substituting a first order Taylor's series approximation for the time delay variation, is (assuming the previous process output is used in its calculation)

$$y_{m8}(n) = e^{-T_s/T_m} y_1(n-1) - \frac{K_m(g_p - g_m)T_s}{T_m} u(n-g_m) - K_m(e^{-T_s/T_m} - 1 - \frac{(g_p - g_m)T_s}{T})u(n-g_m-1) \tag{3.53}$$

Then, from equations (3.12) and (3.53), $e_8(n) = y_1(n) - y_{m8}(n)$ equals

$$\frac{K_m(g_p - g_m)T}{T_m}u(n-g_m) + K_m(e^{-T_s/T_m} - 1 - \frac{(g_p - g_m)T_s}{T_m})u(n-g_m-1) + w(n)$$

$$+K_p(1-e^{-T_s/T_p})u(n-g_p-1) + (e^{-T_s/T_p} - e^{-T_s/T_m})y(n-1) \qquad (3.54)$$

The MSE performance surface, $E[e_g^2(n)]$, may then be calculated (using a procedure similar to that used in equations (3.22) and (3.23)) to be equal to (O'Dwyer (1996m))

$$(e^{-T_s/T_p} - e^{-T_s/T_m})^2 r_{y_1y_1}(0) + 2\frac{K_m^2(g_p - g_m)T_s}{T_m}(e^{-T_s/T_m} - 1 - \frac{(g_p - g_m)T_s}{T_m})r_{uu}(1)$$

$$+[\frac{K_m^2(g_p - g_m)^2T_s^2}{T_m^2} + K_p^2(1-e^{-T_s/T_p})^2 + K_m^2(e^{-T_s/T_m} - 1 - \frac{(g_p - g_m)T_s}{T_m})^2]r_{uu}(0)$$

$$-2K_p(1-e^{-T_s/T_p})[\frac{-K_m(g_p - g_m)T_s}{T_m}r_{uu}(g_p - g_m + 1) - K_m(e^{-T_s/T_m} - 1 - \frac{(g_p - g_m)T_s}{T_m})r_{uu}(g_p - g_m)]$$

$$-2(e^{-T_s/T_p} - e^{-T_s/T_m})[\frac{-K_m(g_p - g_m)T_s}{T_m}r_{uy_1}(g_m - 1) - K_m(e^{-T_s/T_m} - 1 - \frac{(g_p - g_m)T_s}{T_m})r_{uy_1}(g_m)]$$

$$+2(e^{-T_s/T_p} - e^{-T_s/T_m})K_p(1-e^{-T_s/T_p})r_{uy_1}(g_p) + r_{ww}(0) \qquad (3.55)$$

For white noise excitation:

$$r_{uu}(k) = r_{uu}(0) \text{ , } k = 0 \text{ and } r_{uu}(k) = 0 \text{ otherwise.} \qquad (3.56)$$

Also, it may be shown that, for white noise excitation,

$$r_{uy_1}(g_p + n) = (e^{-T_s/T_p})^{n-1}K_p(1-e^{-T_s/T_p})r_{uu}(0), \; n \ge 1 \qquad (3.57)$$

and

$$r_{uy_1}(g_p + n) = 0 \text{ otherwise.} \qquad (3.58)$$

Reason:

$$\begin{aligned} r_{uy_1}(g_p) &= E[y_1(n)u(n-g_p)] \\ &= E[\{e^{-T_s/T_p}y_1(n-1) + K_p(1-e^{-T_s/T_p})u(n-g_p-1)\}u(n-g_p)] \\ &= e^{-T_s/T_p}r_{uy_1}(g_p-1) + K_p(1-e^{-T_s/T_p})r_{uu}(1) \\ &= e^{-T_s/T_p}r_{uy_1}(g_p-1) \end{aligned}$$

Repeated application of this procedure gives equation (3.58).

$$
\begin{aligned}
r_{uy_1}(g_p+1) &= e^{-T_s/T_p} r_{uy_1}(g_p) + K_p(1-e^{-T_s/T_p}) r_{uu}(0) \\
&= K_p(1-e^{-T_s/T_p}) r_{uu}(0) \\
r_{uy_1}(g_p+2) &= e^{-T_s/T_p} r_{uy_1}(g_p+1) + K_p(1-e^{-T_s/T_p}) r_{uu}(1) \\
&= K_p(1-e^{-T_s/T_p}) e^{-T_s/T_p} r_{uu}(0)
\end{aligned}
$$

Repeated application of this procedure gives equation (3.57).

For $g_m = g_p$, the value of the MSE (equation (3.55)) equals

$$
MSE_{opt} = (e^{-T_s/T_p} - e^{-T_s/T_m})^2 r_{y_1y_1}(0) + [K_p(1-e^{-T_s/T_p}) - K_m(1-e^{-T_s/T_m})]^2 r_{uu}(0) + r_{ww}(0) \tag{3.59}
$$

By comparing the amplitudes of the individual terms in equations (3.55) and (3.59), it may be shown that $E[e_8^2(n)] > MSE_{opt}$ for

(a) $g_p > g_m$ (for all values of the other process and model parameters) and

(b) $g_p < g_m$, provided $K_p/K_m \le T_p / T_m$ and $T_p \ge T_m$ (O'Dwyer (1996m)).

The conditions in (b) above are sufficient, rather than necessary, conditions.

However, if a gradient method is used to determine g_p, then an additional restriction that the MSE function must be unimodal about a minimum value when $g_m = g_p$, is imposed. The unimodality of the MSE function in equation (3.55) may be proved by induction; an outline of the inductive proof (provided in full by O'Dwyer (1996m)) is as follows:

(a) For $g_p > g_m$, it may be proved, using equations (3.55) to (3.59), that the MSE function at $g_m = g_p - 1$ is always greater than the MSE function at $g_m = g_p$. Similarly, using equations (3.55) to (3.58), it may be proved that the MSE function at $g_m = g_p - n - 1$ is always greater than the MSE function at $g_m = g_p - n$.

(b) For $g_p < g_m$, it may be proved, using equations (3.55) to (3.59), that the MSE function at $g_m = g_p + 1$ is always greater than the MSE function at $g_m = g_p$, provided

$$T_p \geq T_m \text{ and } K_p / K_m \leq T_p / T_m \tag{3.60}$$

These are sufficient, rather than necessary conditions. Similarly, using equations (3.55) to (3.58), it may be proved that the MSE function at $g_m = g_p + n + 1$ is greater than the MSE function at $g_m = g_p + n$, provided

$$\frac{K_m T_s}{T_m}[(2n+1)\frac{T_s}{T_m} - (1 - e^{-T_s/T_m})]$$
$$+K_p(1 - e^{-T_s/T_p})^2 e^{-(n-2)T_s/T_p}(e^{-T_s/T_p} - e^{-T_s/T_m})\{(1 - e^{-T_s/T_m})e^{-T_s/T_p} - \frac{T_s}{T_m}[(n+1)e^{-T_s/T_p} - n)]\} > 0 \tag{3.61}$$

This is a necessary condition. □

This theorem indicates that if K_p and T_p are unknown, then convergence of the model time delay index to the process time delay index may only be completely guaranteed if the value of the model time delay index is always less than or equal to the process time delay index. The nature of typical MSE functions versus model time delay index is shown by Figures 3.22 and 3.23. In Figure 3.22, $K_p = 2.0$, $K_m = 1.0$, $T_p = 2.0$ seconds and $T_m = 1.0$ seconds, so that the conditions in equations (3.60) and (3.61) are fulfilled. In Figure 3.23, $K_p = 2.0$, $K_m = 1.0$, $T_p = 0.7$ seconds and $T_m = 1.0$ seconds, so that the conditions in equation (3.60) are violated. In both simulations, $g_p = 30$. The normalised MSE (equal to the MSE divided by $r_{uu}(0)$) versus model time delay index is plotted in both cases, with $r_{ww}(0)$ put to zero. As expected, the normalised MSE function is unimodal with respect to the model time delay index, with a minimum value at $g_p = g_m$ when the conditions in equations (3.60) and (3.61) are fulfilled; when the conditions given by equation (3.60) are violated, the MSE curve is still unimodal, but it has a minimum value when $g_p \neq g_m$.

Figure 3.22: Normalised MSE vs. time delay index (conditions in equation (3.60) met)

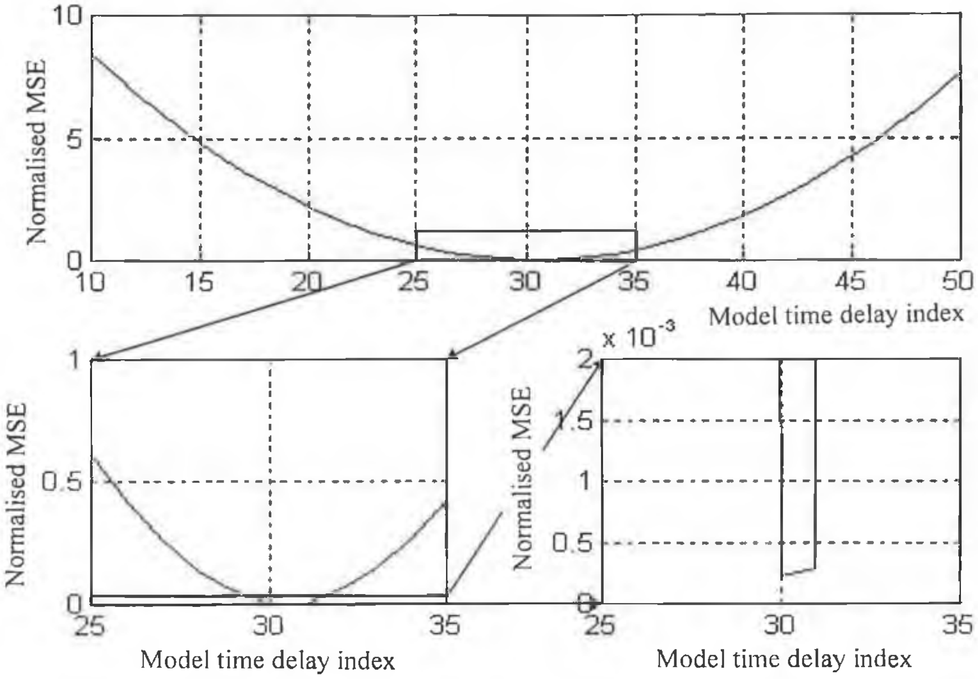
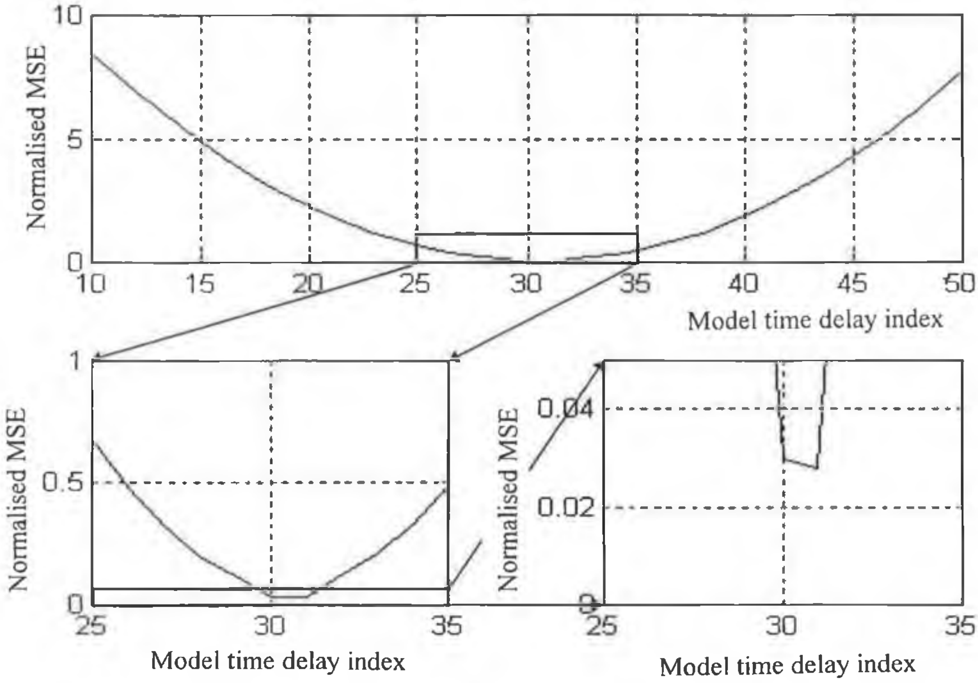
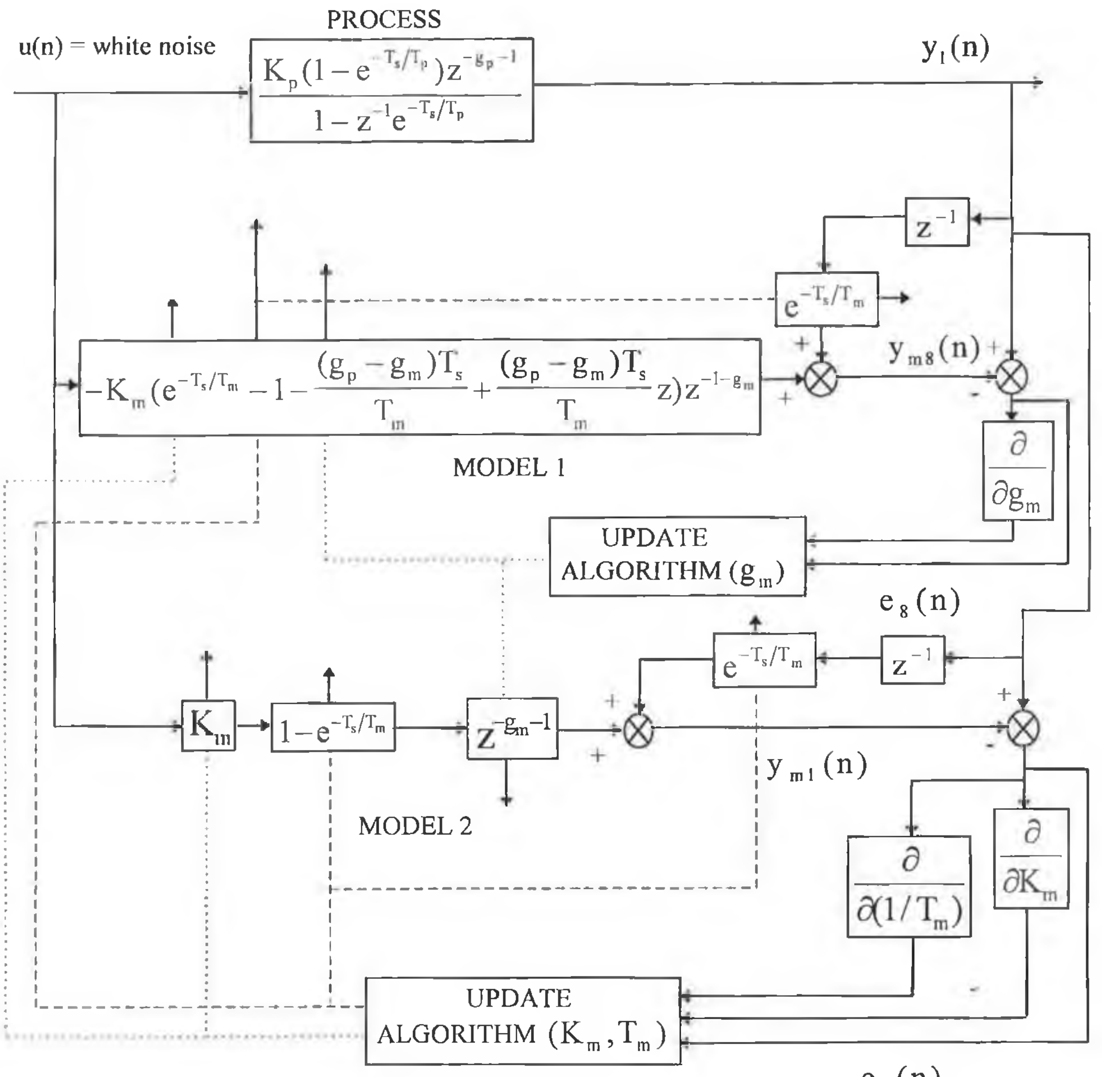


Figure 3.23: Normalised MSE vs. time delay index (conditions in equation (3.60) violated)



Considering equations (3.12), (3.13), (3.14), (3.53) and (3.54), a block diagram representation of the gradient method to update both the model parameters and the model time delay index is as shown in Figure 3.24. The non-delay model parameters are updated based on a FOLPD process model.

Figure 3.24: Updating the full parameter set - Case 1



One set of representative simulation results corresponding to Theorem 3.5 are given in Figures 3.25a-3.25d and 3.26a-3.26d, with the parameters and the process minus model output plotted against sample number. At the beginning, the starting values of the gain, time constant and time delay index for both the process and model are equal; a step change was then made to the process parameter values. The parameter values are taken as $K_p = 2.0$, $K_m = 1.0$, $T_p = 2.0$ seconds and $T_m = 1.0$ seconds, so that the conditions for the unimodality of the MSE function (given by equations (3.60) and (3.61)) are fulfilled; Figure 3.22 shows the corresponding MSE function.

Figure 3.25a: Gain estimate

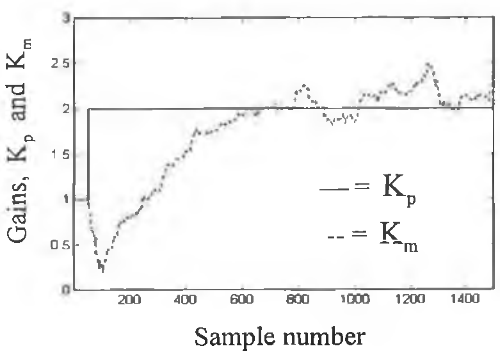


Figure 3.25b: Time constant estimate

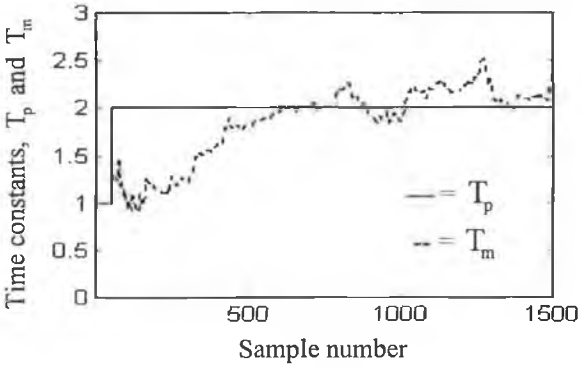


Figure 3.25c: Time delay index estimate

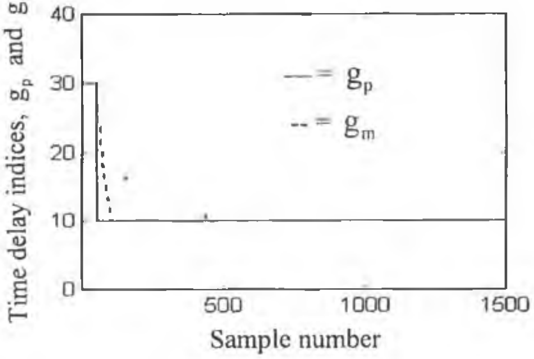


Figure 3.25d: $e_1(n)$

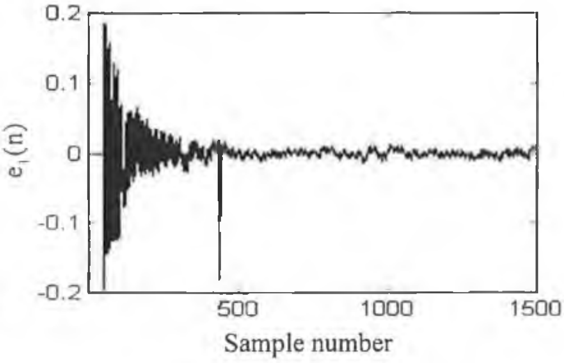


Figure 3.26a: Gain Estimate

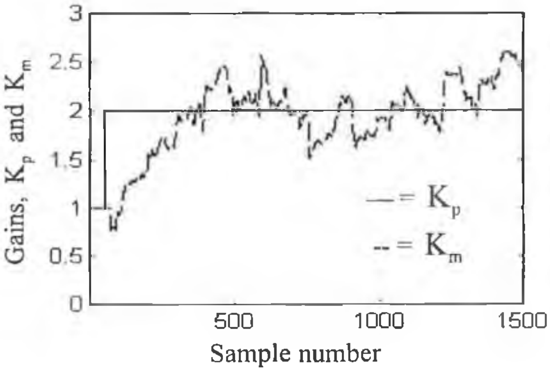


Figure 3.26b: Time constant Estimate

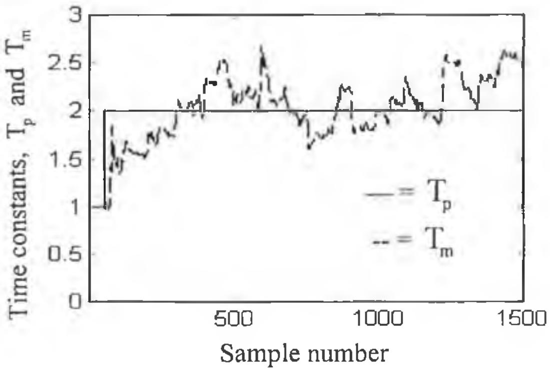


Figure 3.26c: Time delay index Estimate

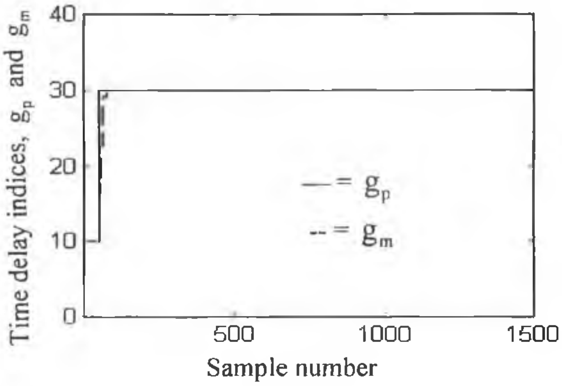
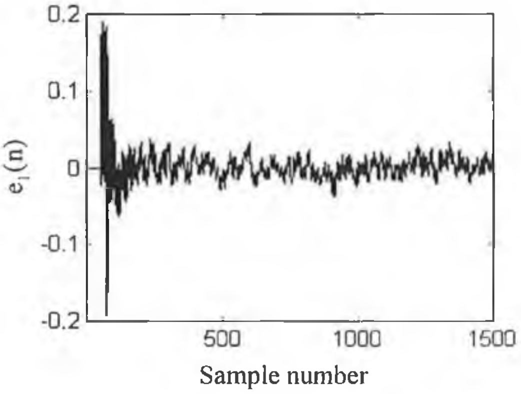


Figure 3.26d: $e_1(n)$



The simulation conditions for updating the model time delay index are identical to those in Section 3.4.1 (with the excitation signal assumed to be white noise). The model gain and time constant were updated in a similar manner, with these parameter estimates filtered by a low pass filter. A lower limit was also put on the model time constant of 0.1 times the starting value of the model time constant (it was determined in simulation that this lower limit is necessary, as it is the reciprocal of the time constant that is being updated). A lower limit of zero is placed on the model gain and time delay index value.

Fast convergence to the process parameter values is seen for a relatively low level of coloured measurement noise, for both cases of interest (i.e. when the starting value of the model time delay index is in turn, greater than and less than the process time delay index). This is as expected from Theorem 3.5.

If the level of coloured measurement noise is greater than is taken in the simulations corresponding to Figures 3.25a-3.25d and Figures 3.26a-3.26d, reasonable (if noisy) convergence to the correct value of the process time delay index is observed (see Figures 3.27a and 3.27b).

Figure 3.27a: Time delay index estimate

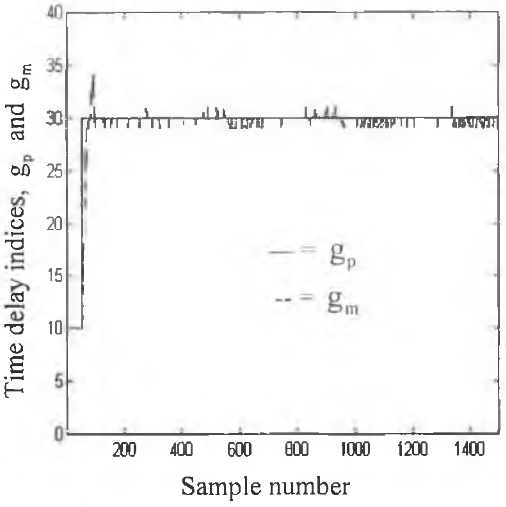
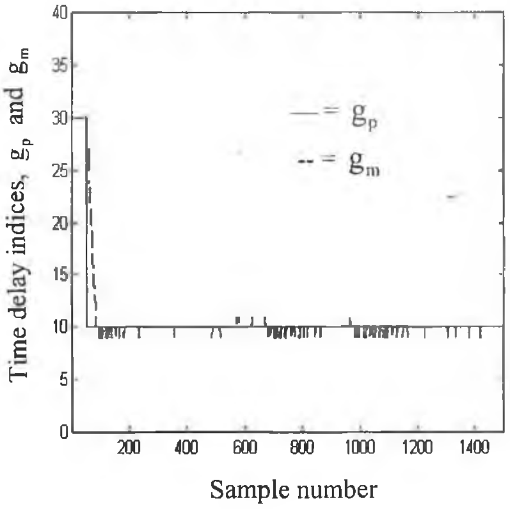


Figure 3.27b: Time delay index estimate



The simulation results discussed are interesting because they show that the model gain and time constant values converge to the process gain and time constant values, in the presence of coloured measurement noise (which is predicted from Theorem 3.1). Generally speaking, the model time delay index must converge to the process time delay index before convergence of the other parameters is observed. This was also explored theoretically in Theorem 3.1. Simulation work has revealed that there does seem to be a (relatively low) level of coloured measurement noise above which the model gain and time constant estimates do not converge to the process gain and time constant values, because of the noisy convergence of the model time delay index estimate; a greater level of filtering on this parameter could be helpful. Of course, as in Section 3.4.1, the procedures outlined depend on *a priori* knowledge of the process time delay index, g_p.

In Theorem 3.5, the previous process output is used to calculate the new model output. The previous model output could be used to calculate the new model output; the model difference equation under these circumstances is (when a first order Taylor's series is used to approximate the time delay variation)

$$y_{m9}(n) = e^{-T_s/T_m} y_{m9}(n-1) - \frac{K_m(g_p - g_m)T_s}{T_m} u(n - g_m) - K_m(e^{-T_s/T_m} - 1 - \frac{(g_p - g_m)T_s}{T_m}) u(n - g_m - 1) \tag{3.62}$$

Supplementary simulation results confirm that convergence of the model parameters to the process parameters was achieved, when a model corresponding to equation (3.62) was used to determine the gradient algorithm, under the circumstances discussed in this section of the chapter.

The unimodality of the MSE function versus model time delay index for this case may also be proved by calculating the MSE function using the method of Widrow and Stearns (1985). If the residue theorem is used to calculate the required closed curve integral, it may be determined that the MSE performance surface is quadratic in r when the MSE is calculated based on the model in equation (3.62), and is thus unimodal in this variable i.e. if, from equations (3.12) and (3.62), $e_9(n) = y_1(n) - y_{m9}(n)$, then, following the procedure given in equations (3.15) and (3.16), it may be proven that (O'Dwyer (1996m))

$$E[e_9^2(n)] = A\left(\frac{r}{T_m}\right)^2 + B\left(\frac{r}{T_m}\right) + C \tag{3.63}$$

with
$$A = \frac{2K_m^2(1 - e^{-T_s/T_m})}{1 - e^{-2T_s/T_m}} \tag{3.64}$$

$$B = \frac{2K_m^{\,2}(1-e^{-T_s/T_m})^2}{1-e^{-2T_s/T_m}} - \frac{2K_mK_p(1-e^{-T_s/T_m})(1-e^{-T_s/T_p})}{1-e^{-T_s/T_m}e^{-T_s/T_p}} \tag{3.65}$$

and

$$C = \frac{K_m^{\,2}(1-e^{-T_s/T_m})^2}{1-e^{-2T_s/T_m}} - \frac{2K_mK_p(1-e^{-T_s/T_m})(1-e^{-T_s/T_p})}{1-e^{-T_s/T_m}e^{-T_s/T_p}} + \frac{K_p^{\,2}(1-e^{-T_s/T_p})^2}{1-e^{-2T_s/T_p}} \tag{3.66}$$

If, in addition, $K_p = K_m = K$ and $T_p = T_m = T$, then the MSE function equals $E[e_7^{\,2}(n)]$ (with $e_7(n)$ given by equation (3.43)), which may be calculated, by the use of equations (3.63) to (3.66), to be

$$E[e_7^{\,2}(n)] = \frac{2K^2(1-e^{-T_s/T})}{1-e^{-2T_s/T}}\left(\frac{r}{T}\right)^2 \tag{3.67}$$

This function is unimodal, and is minimised when $r = 0$ (i.e. when $g_p = g_m$).

3.5.2 Convergence of the full parameter set - Case 2

In this case, the procedures defined do not depend on *a priori* knowledge of the time delay index, g_p (as in the procedures defined in Section 3.4.2).

3.5.2.1 The time delay as an integer multiple of the sample period - white noise input

Theorem 3.6: For a first order discrete stable system of unknown parameters, the MPE performance surface versus model time delay index is minimised when the model time delay index equals the process time delay index, under the following conditions:

(a) The time delay variation is approximated by a first order Taylor's series approximation

(b) The measurement noise is uncorrelated with the process input and output

(c) The resolution on the process time delay is assumed to be equal to one sample period

(d) The error is calculated based on using a FOLPD process model; the partial derivative of the error with respect to the time delay variation is calculated based on the first order Taylor's series approximation for the time delay variation

(e) The conditions provided in the theorem are observed on the model parameters and

(f) The input to the model and the process is assumed to be a white noise signal.

Proof: The process difference equation, $y_1(n)$, is given by equation (3.12). The corresponding model difference equation, $y_{m1}(n)$, is given by equation (3.13) (assuming the previous process output is used in its calculation). The model difference equation for calculating the partial derivative of the error with respect to the time delay variation, $y_{m8}(n)$, is given by equation (3.53) (assuming that the previous process output is used in its calculation). The error, $e_1(n) = y_1(n) - y_{m1}(n)$, is given by equation (3.14) and the error, $e_8(n) = y_1(n) - y_{m8}(n)$, is given by equation (3.54). The MPE performance surface, $E[e_1(n)e_8(n)]$, may then be calculated, using a procedure similar to that used in equations (3.22) and (3.23), to be equal to (O'Dwyer (1996m))

$$\begin{aligned}
&(e^{-T_s/T_p} - e^{-T_s/T_m})^2 r_{y_1y_1}(0) \\
&\quad +[K_p{}^2(1-e^{-T_s/T_p})^2 + K_m{}^2(1-e^{-T_s/T_m})(1-e^{-T_s/T_m} + \frac{(g_p-g_m)T_s}{T_m})]r_{uu}(0) \\
&+K_pK_m(1-e^{-T_s/T_p})\{\frac{(g_p-g_m)T_s}{T_m} r_{uu}(g_p-g_m+1) - [2(1-e^{-T_s/T_m}) + \frac{(g_p-g_m)T_s}{T_m}]r_{uu}(g_p-g_m)\} \\
&+K_m(e^{-T_s/T_p} - e^{-T_s/T_m})\{\frac{(g_p-g_m)T_s}{T_m}[r_{uy_1}(g_m-1) - r_{uy_1}(g_m)] - 2(1-e^{-T_s/T_m})r_{uy_1}(g_m)\} + r_{ww}(0) \\
&\qquad\qquad -K_m{}^2(1-e^{-T_s/T_m})\frac{(g_p-g_m)T_s}{T_m} r_{uu}(1)
\end{aligned} \tag{3.68}$$

with $r_{uu}(n)$ and $r_{uy_1}(n)$ provided in equations (3.56), (3.57) and (3.58) respectively. For white noise excitation, at $g_m = g_p$, the value of the MPE equals (using equations (3.56) to (3.58) and equation (3.68))

$$MPE_{opt} = (e^{-T_s/T_p} - e^{-T_s/T_m})^2 r_{y_1y_1}(0) + [K_p(1-e^{-T_s/T_p}) - K_m(1-e^{-T_s/T_m})]^2 r_{uu}(0) + r_{ww}(0) \tag{3.69}$$

By comparing the amplitudes of the individual terms in equations (3.68) and (3.69), it may be shown that $E[e_1(n)e_8(n)] > MPE_{opt}$ for

(a) $g_p > g_m$ (for all values of process and model parameters)

and

(b) $g_p < g_m$, provided $K_p/K_m \geq (g_m - g_p)/2$ and $T_m \geq T_p$ (O'Dwyer (1996m)).

The conditions in (b) are sufficient, rather than necessary conditions.

However, if a gradient method is used to determine g_p, then an additional restriction that the MPE function must be unimodal with a minimum MPE value occurring at $g_m = g_p$, is imposed. The conditions for unimodality may be proved by induction; an outline of the inductive proof (provided in full by O'Dwyer (1996m)) is as follows:

(a) $g_p > g_m$: It may be proved, using equations (3.56) to (3.58) and equations (3.68) and (3.69), that the MPE function at $g_m = g_p - 1$ is greater than the MPE function at $g_m = g_p$. Similarly, using equations (3.56) to (3.58) and equation (3.68), it may be proved that the MPE function at $g_m = g_p - n - 1$ is always greater than the MPE function at $g_m = g_p - n$.

(b) $g_p < g_m$: It may be proved, using equations (3.56) to (3.58) and equations (3.68) and (3.69), that the MPE function at $g_m = g_p + 1$ is greater than the MPE function at $g_m = g_p$, provided that the following sufficient conditions are obeyed:

$$(1 - T_s/T_m)K_p(1 - e^{-T_s/T_p}) > K_m(1 - e^{-T_s/T_m}) \text{ and } T_m > T_p \tag{3.70}$$

The nature of the MPE function means that for a full inductive proof, it is necessary to prove that the MPE function at $g_m = g_p + 2$ is greater than the MPE function at $g_m = g_p + 1$ (this is because the MPE function in equation (3.68) depends on $r_{uu}(g_p - g_m + 1)$). A necessary condition for this to be true, using equations (3.56) to (3.58) and equation (3.68), is if

$$\frac{T_s}{T_m}[K_p(1 - e^{-T_s/T_p}) - K_m(1 - e^{-T_s/T_m})] >$$

$$(e^{-T_s/T_m} - e^{-T_s/T_p})K_p(1 - e^{-T_s/T_p})[2(1 - e^{-T_s/T_m} - \frac{T_s}{T_m})(1 - e^{-T_s/T_p}) - \frac{T_s}{T_m}] \qquad (3.71)$$

Similarly, it may be proved, using equations (3.56) to (3.58) and equation (3.68), that the MPE function at $g_m = g_p + n + 1$ is greater than the MPE function at $g_m = g_p + n$, provided that

$$\frac{K_m(1 - e^{-T_s/T_m})\frac{T_s}{T_m}}{K_p(1 - e^{-T_s/T_p})e^{-(n-2)T_s/T_p}(e^{-T_s/T_p} - e^{-T_s/T_m})} <$$
$$\frac{T_s}{T_m}[(n+1)e^{-2T_s/T_p} - (2n+1)e^{-T_s/T_p} + n] + 2(1 - e^{-T_s/T_m})(1 - e^{-T_s/T_p})e^{-T_s/T_p} \qquad (3.72)$$

This is a necessary condition. □

The theorem indicates that if K_p and T_p are unknown, then convergence of the model time delay index to the process time delay index may only be completely guaranteed if the value of the model time delay index is always less than or equal to the process time delay index. The behaviour of the MPE function (given by equation (3.68)) versus time delay index is confirmed, in representative simulation results, by Figures 3.28 and 3.29. In Figure 3.28, $K_p = 2.0$, $K_m = 1.0$, $T_p = 0.7$ seconds and $T_m = 1.0$ seconds so that the conditions given in equations (3.70) and (3.71) (but not (3.72)) are fulfilled; in Figure 3.29, $K_p = 2.0$, $K_m = 3.0$, $T_p = 0.7$ seconds and $T_m = 0.5$ seconds, so that none of the conditions in equations (3.70), (3.71) or (3.72) are fulfilled (the former conditions are identical to those used to calculate the normalised MSE curve in Figure 3.23). The normalised MPE (equal to the MPE divided by $r_{uu}(0)$) is plotted versus time delay index in both cases, with $r_{ww}(0)$ put to zero and $g_p = 30$. The excitation signal in both cases is a white noise signal. The results are as expected from the theorem.

Figure 3.28: Normalised MPE vs. time delay index - white noise excitation

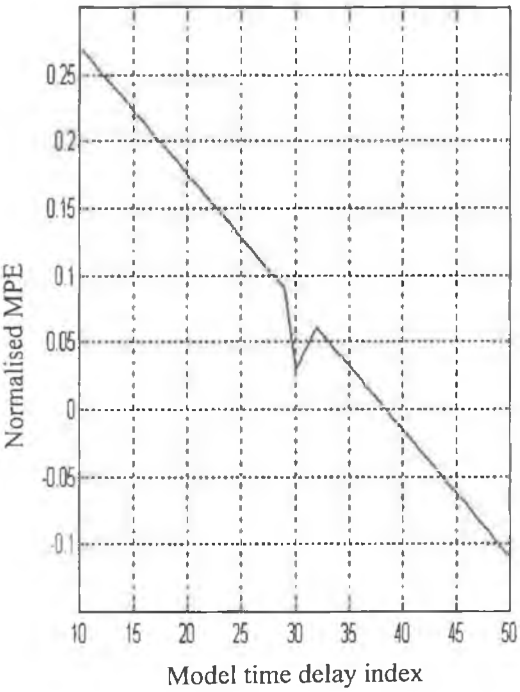
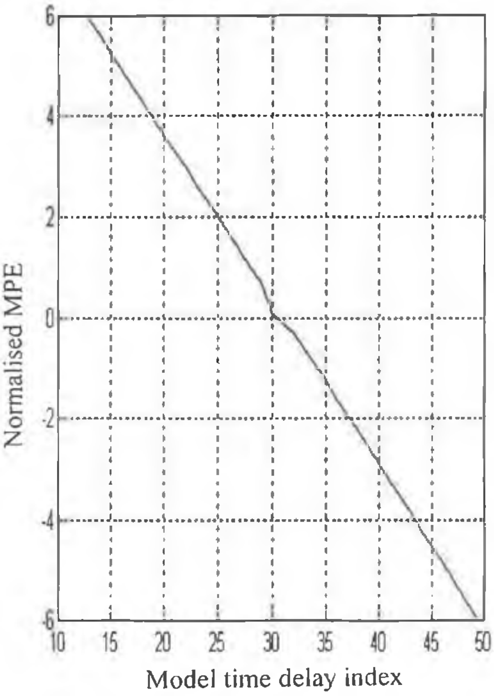


Figure 3.29: Normalised MPE vs. time delay index - white noise excitation



Considering equations (3.12), (3.13), (3.14), (3.53) and (3.54), a block diagram representation of the gradient method to update both the model parameters and the model time delay index may be represented as shown in Figure 3.30; as in Section 3.5.1, the non-delay model parameters are updated based on a FOLPD process model.

A representative simulation result corresponding to Theorem 3.6 is given in Figures 3.31a-3.31d, with the parameters and the process minus model output plotted against sample number. The simulation conditions for updating the time delay are identical to those in Section 3.5.1, though the process parameter variations considered are different. It was also found necessary to limit the variation of the non-time delay model parameters; for the simulations taken, $0.5 < K_m < 3.0$ and 0.5 seconds $< T_m < 3.0$ seconds were the limits. The normalised MPE curve corresponding to these simulation results is given by Figure 3.28.

Figure 3.30: Updating of the full parameter set - Case 2

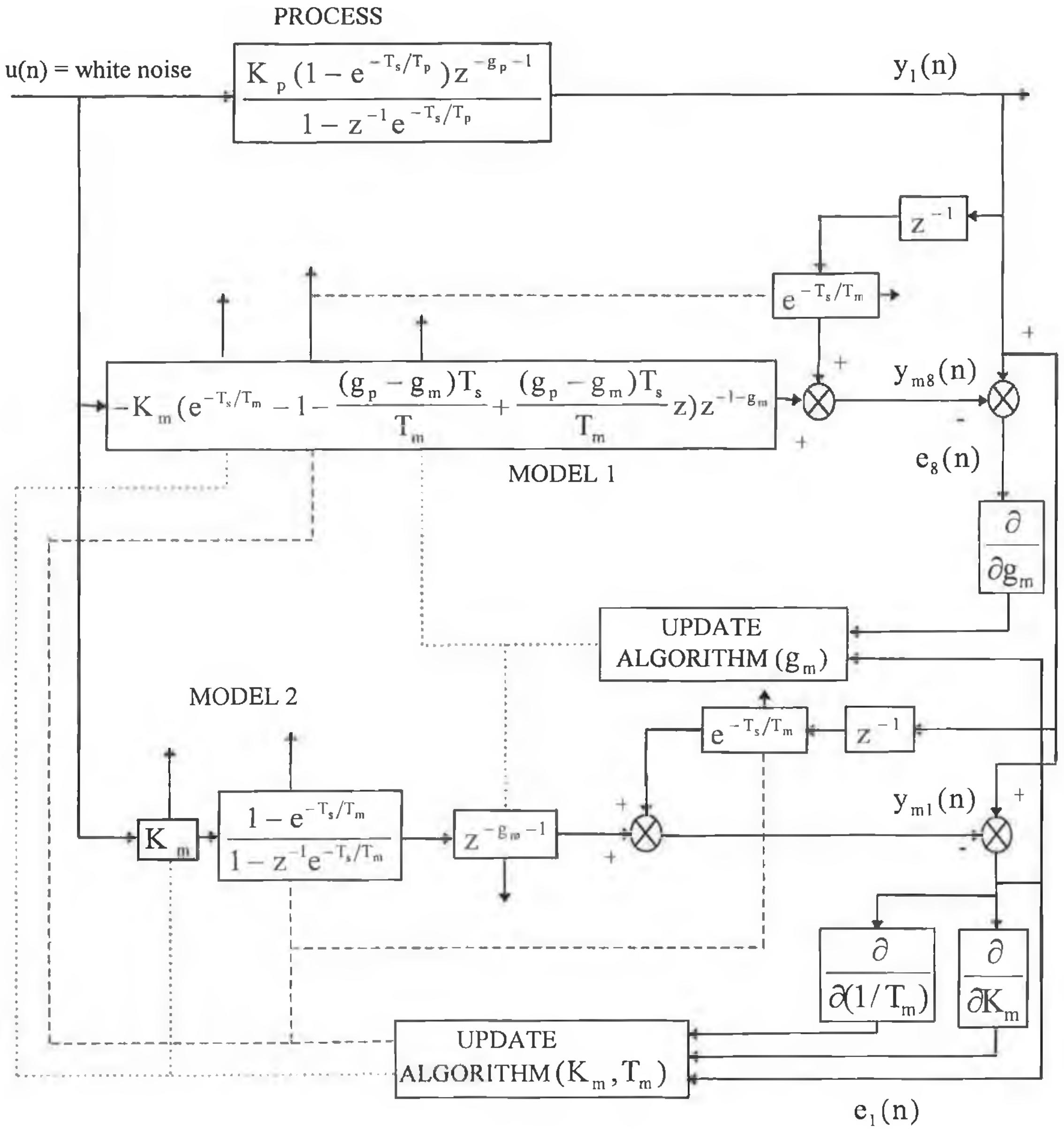


Figure 3.31a: Gain estimate

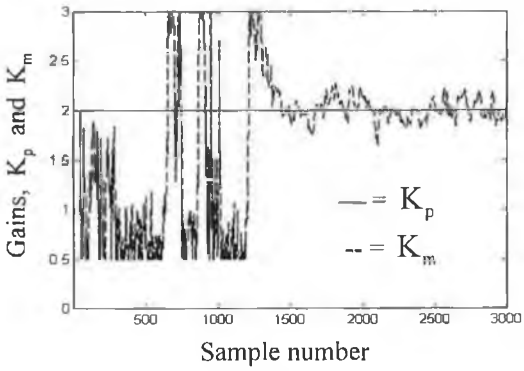


Figure 3.31b: Time constant estimate

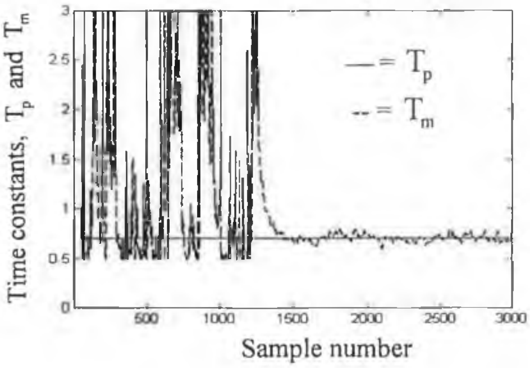


Figure 3.31c: Time delay index estimate

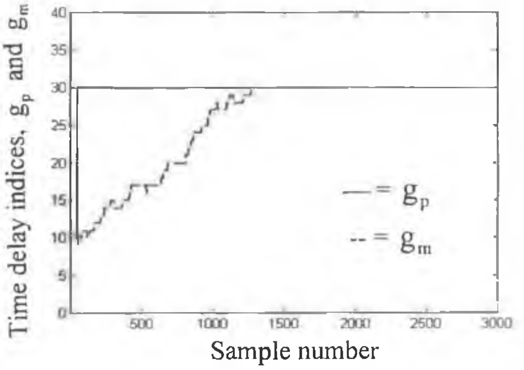
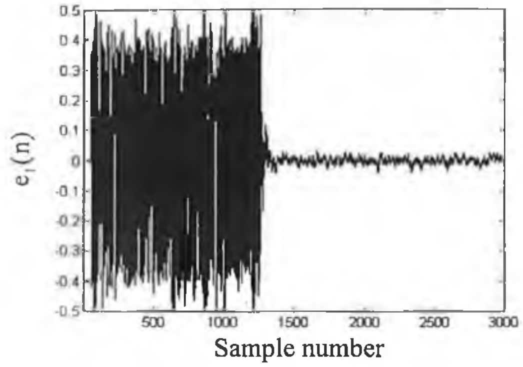


Figure 3.31d: $e_1(n)$



These results conform with Theorem 3.6.

In this theorem, the previous process output is used to calculate the new model output. The previous model output could also be used to calculate the new model output, with the model difference equation being given by equation (3.62). Simulation results confirm that convergence of the model parameters to the process parameters also results in this case, under the circumstances discussed in this section of the chapter.

3.5.2.2 The time delay as a real multiple of the sample period - white noise input

Theorem 3.6 has dealt with the estimation of process time delays that are integer multiples of the sample period. For the estimation of process time delays that are real multiples of the sample period, then the difference equation of a FOLPD process is (O'Dwyer (1996e))

$$y_4(n) = e^{-T_s/T_p} y_4(n-1) + K_p(1 - e^{g_b T_s/T_p})u(n-g_p) + K_p(e^{g_b T_s/T_p} - e^{-T_s/T_p})u(n-g_p-1) + w(n) \quad (3.73)$$

The corresponding model difference equation (assuming the previous process output is used in its calculation) is

$$y_{m10}(n) = e^{-T_s/T_m} y_4(n-1) + K_m(1 - e^{g_a T_s/T_m})u(n-g_m) + K_m(e^{g_a T_s/T_m} - e^{-T_s/T_m})u(n-g_m-1) \quad (3.74)$$

The model difference equation for calculating the partial derivative of the error with respect to the time delay variation (and assuming that the previous process output is used in its calculation) is

$$y_{m11}(n) = e^{-T_s/T_m} y_4(n-1) - \frac{K_m(g_p - g_m + g_b - g_a)T_s}{T_m} u(n-g_m)$$

$$-K_m[e^{-T_s/T_m} - 1 - \frac{(g_p - g_m + g_b - g_a)T_s}{T_m}]u(n-g_m-1) \qquad (3.75)$$

This equation may be deduced from equation (3.53). Therefore, from equations (3.73) and (3.74),

$$e_{10}(n) = y_4(n) - y_{m10}(n) = (e^{-T_s/T_p} - e^{-T_s/T_m})y_4(n-1) + K_p(1 - e^{g_b T_s/T_p})u(n-g_p)$$

$$+K_p\left(e^{g_b T_s/T_p} - e^{-T_s/T_p}\right)u(n-g_p-1)$$

$$-K_m(1 - e^{g_a T_s/T_m})u(n-g_m) - K_m(e^{g_a T_s/T_m} - e^{-T_s/T_m})u(n-g_m-1) + w(n) \qquad (3.76)$$

and, from equations (3.73) and (3.75),

$$e_{11}(n) = y_4(n) - y_{m11}(n) = (e^{-T_s/T_p} - e^{-T_s/T_m})y_4(n-1) + K_p(1 - e^{g_b T_s/T_p})u(n-g_p)$$

$$+K_p(e^{g_b T_s/T_p} - e^{-T_s/T_p})u(n-g_p-1)$$

$$+\frac{K_m T_s(g_p - g_m + g_b - g_a)}{T_m} u(n-g_m)$$

$$+K_m[e^{-T_s/T_m} - 1 - \frac{T_s(g_p - g_m + g_b - g_a)}{T_m}]u(n-g_m-1) + w(n) \qquad (3.77)$$

The MPE performance surface, $E[e_{10}(n)e_{11}(n)]$, may then be calculated, using a procedure similar to that used in equations (3.22) and (3.23), to be equal to (O'Dwyer (1996m))

$$(e^{-T_s/T_p} - e^{-T_s/T_m})^2 r_{y_4 y_4}(0) + 2(e^{-T_s/T_p} - e^{-T_s/T_m})K_p(1 - e^{g_b T_s/T_p})r_{uy_4}(g_p - 1)$$

$$+2(e^{-T_s/T_p} - e^{-T_s/T_m})K_p(e^{g_b T_s/T_p} - e^{-T_s/T_p})r_{uy_4}(g_p)$$

$$
\begin{gathered}
+(e^{-T_s/T_p}-e^{-T_s/T_m})[\frac{K_m T_s(g_p-g_m+g_b-g_a)}{T_m}-K_m(1-e^{g_a T_s/T_m})]r_{uy_4}(g_m-1) \\
+(e^{-T_s/T_p}-e^{-T_s/T_m})[K_m\{e^{-T_s/T_m}-1-\frac{T_s(g_p-g_m+g_b-g_a)}{T_m}-K_m(e^{g_a T_s/T_m}-e^{-T_s/T_m})]r_{uy_4}(g_m)+ \\
[K_p{}^2(1-e^{g_b T_s/T_p})^2+K_p{}^2(e^{g_b T_s/T_p}-e^{-T_s/T_p})^2-K_m{}^2(1-e^{g_a T_s/T_m})\frac{T_s}{T_m}(g_p-g_m+g_b-g_a)]r_{uu}(0) \\
-K_m{}^2(e^{g_a T_s/T_m}-e^{-T_s/T_m})[e^{-T_s/T_m}-1-\frac{T_s}{T_m}(g_p-g_m+g_b-g_a)]r_{uu}(0)\ + \\
[K_p(1-e^{g_b T_s/T_p})[\frac{K_m(g_p-g_m+g_b-g_a)T_s}{T_m}-K_m(1-e^{g_a T_s/T})]r_{uu}(g_p-g_m)+ \\
K_p(e^{g_b T_s/T_p}-e^{-T_s/T_p})[K_m\{e^{-T_s/T_m}-1-\frac{(g_p-g_m+g_b-g_a)T_s}{T_m}\}-K_m(e^{g_a T_s/T_m}-e^{-T_s/T_m})]r_{uu}(g_p-g_m) \\
+K_p(1-e^{g_b T_s/T_p})[K_m\{e^{-T_s/T_m}-1-\frac{(g_p-g_m+g_b-g_a)T_s}{T_m}\}-K_m(-e^{-T_s/T_m}+e^{g_a T_s/T_m})]r_{uu}(g_p-g_m-1) \\
+K_p(e^{g_b T_s/T_p}-e^{-T_s/T_p})[\frac{K_m(g_p-g_m+g_b-g_a)T_s}{T_m}-K_m(1-e^{g_a T_s/T_m})]r_{uu}(g_p-g_m+1)+r_{ww}(0) \\
+[2K_p{}^2(1-e^{g_b T_s/T_p})(e^{g_b T_s/T_p}-e^{-T_s/T_p})-\frac{K_m{}^2}{T_m}(2e^{g_a T_s/T_m}-1-e^{-T_s/T_m})T_s(g_p-g_m+g_b-g_a)]r_{uu}(1) \\
+K_m{}^2(1-e^{-T_s/T_m})(1-e^{g_a T_s/T_m})r_{uu}(1) \qquad (3.78)
\end{gathered}
$$

with $r_{y_4y_4}(n)$ being the autocorrelation function of $y_4(n)$. The behaviour of the MPE function versus model time delay index is given, in representative simulation results, by Figures 3.32 and 3.33. For these simulations, $K_p = 2.0$, $K_m = 1.0$, $T_p = 0.7$ seconds, $T_m = 1.0$ seconds and $g_p = 5$, with the time delay taken in intervals of 0.01 times the sample period. The normalised MPE (equal to the MPE divided by $r_{uu}(0)$) is plotted versus model time delay index for $g_b = 0.0$ and $g_b = 0.5$. In Figure 3.32 the normalised MPE calculated from equation (3.68) is superimposed on the plots for comparison purposes. For all simulations, $r_{ww}(0)$ is put to zero. Clearly, the MPE function is multimodal with respect to time delay, when the time delay is a real multiple of the sample period (as in Section 3.4.2.2). The estimation of the real value of the time delay is therefore impossible using gradient methods.

Figure 3.32: Normalised MPE vs. time delay index - white noise excitation - $g_b = 0.0$

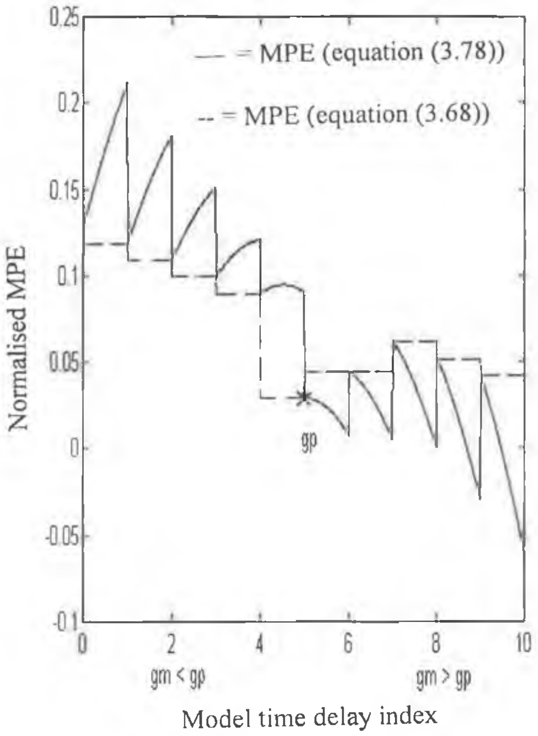
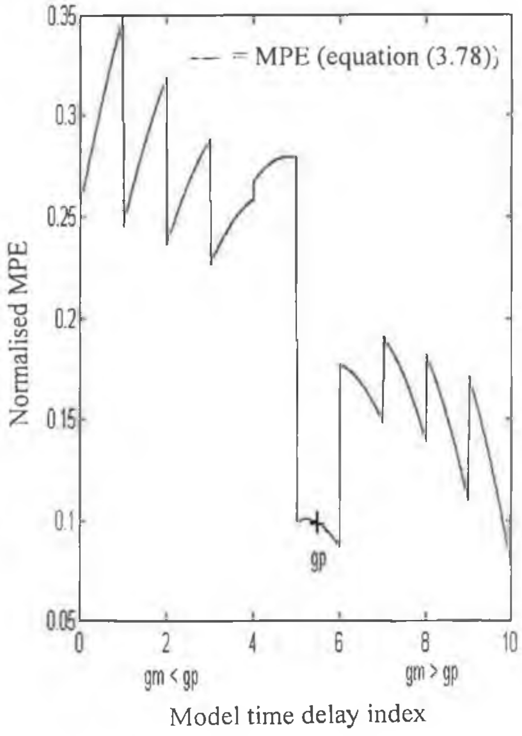


Figure 3.33: Normalised MPE vs. time delay index - white noise excitation - $g_b = 0.5$



3.5.2.3 Conclusions

Overall, the gradient method will allow the estimation of process parameters, for white noise excitation, provided that the process time delay is an integer multiple of the sample period, and provided that the process time delay is always greater than the model time delay, as the model time delay converges to the process time delay (or, if the process time delay is less than the model time delay, the conditions in equations (3.70), (3.71) and (3.72) must be fulfilled). This conclusion is broadly analogous to the conclusion in Section 3.4.2.3. The use of a non-gradient method of estimating the process time delay index that involves determining the minimum positive MPE value is also viable, at least when the process is excited by white noise. Unfortunately, it is not possible to estimate the process parameters if the process time delay is not an integer multiple of the sample period, using the gradient method. On the other hand, the indications from Figures 3.32 and 3.33 are that a non-gradient method of estimating the process time delay, based on calculating the minimum positive value of the MPE surface, may allow the estimation of the correct value of process time delay, when the process time delay is not an integer multiple of the sample period (at least for $g_p > g_m$). It is interesting that this latter method could perhaps be used to estimate the process time delay when $g_p < g_m$, if g_m is close to g_p (at least in the simulations taken).

3.5.2.4 The time delay as an integer multiple of the sample period - square wave input

Theorem 3.7: For a first order discrete stable system of unknown parameters, the MPE performance surface versus model time delay index is unimodal, with a minimum value of the MPE occurring when the model time delay index equals the process time delay index, under the following conditions:

(a) The time delay variation is approximated by a first order Taylor's series approximation

(b) The measurement noise is uncorrelated with the process input and output

(c) The resolution on the process time delay is assumed to be equal to one sample period

(d) The error is calculated based on using a FOLPD process model; the partial derivative of the error, with respect to the time delay variation, is calculated based on the first order Taylor's series approximation for the time delay variation

(e) The conditions provided in the theorem are observed on the model parameters and

(f) The excitation signal input is a square wave with a half period greater than the maximum possible process time delay.

Proof: The process difference equation, $y_1(n)$, is given by equation (3.12). The corresponding model difference equation, $y_{m1}(n)$, is given by equation (3.13) (assuming the previous process output is used in its calculation). The model difference equation for calculating the partial derivative of the error with respect to the time delay variation, $y_{m8}(n)$, is given by equation (3.53) (assuming that the previous process output is used in its calculation). The error, $e_1(n) = y_1(n) - y_{m1}(n)$, is given by equation (3.14) and the error, $e_8(n) = y_1(n) - y_{m8}(n)$ is given by equation (3.54). The MPE performance surface, $E[e_1(n)e_8(n)]$, is given by equation (3.68).

For a square wave excitation signal of amplitude ± 1: $-1 \le r_{uu}(k) \le 1$

It may be shown that, for square wave excitation,

$$r_{uy_1}(g_m + 1) = K_p(1 - e^{-T_s/T_p}) \\ [(e^{-T_s/T_p})^{g_m} r_{uu}(0 - g_p) + (e^{-T_s/T_p})^{g_m - 1} r_{uu}(1 - g_p) + \ldots\ldots + r_{uu}(g_m - g_p)] \tag{3.79}$$

Reason: $r_{uy_1}(1) = e^{-T_s/T_p} r_{uy_1}(0) + K_p(1 - e^{-T_s/T_p}) r_{uu}(0 - g_p)$

$$= K_p(1 - e^{-T_s/T_p}) r_{uu}(0 - g_p)$$

$$r_{uy_1}(2) = e^{-T_s/T_p} r_{uy_1}(1) + K_p(1 - e^{-T_s/T_p}) r_{uu}(1 - g_p)$$

$$= K_p(1 - e^{-T_s/T_p})[e^{-T_s/T_p} r_{uu}(0 - g_p) + r_{uu}(1 - g_p)]$$

$$r_{uy_1}(3) = e^{-T_s/T_p} r_{uy_1}(2) + K_p(1 - e^{-T_s/T_p}) r_{uu}(2 - g_p)$$

$$= K_p(1 - e^{-T_s/T_p})[(e^{-T_s/T_p})^2 r_{uu}(0 - g_p) + e^{-T_s/T_p} r_{uu}(1 - g_p) + r_{uu}(2 - g_p)]$$

Repeated application of this procedure gives equation (3.79).

For $g_m = g_p$, the value of $E[e_1(n)e_8(n)]$ equals MPE_{opt}, which is given by (using equations (3.68) and (3.79))

$$[K_p(1 - e^{-T_s/T_p}) - K_m(1 - e^{-T_s/T_m})]^2 r_{uu}(0) + 2(e^{-T_s/T_p} - e^{-T_s/T_m}) K_p(1 - e^{-T_s/T_p}) r_{uy_1}(g_p) \\ -(e^{-T_s/T_p} - e^{-T_s/T_m}) 2K_m K_p (1 - e^{-T_s/T_m})(1 - e^{-T_s/T_p})[(e^{-T_s/T_m})^{g_m - 1} r_{uu}(0 - g_p) + \ldots\ldots + r_{uu}(g_m - g_p - 1)] \\ +(e^{-T_s/T_p} - e^{-T_s/T_m})^2 r_{y_1y_1}(0) + r_{ww}(0) \tag{3.80}$$

By comparing the amplitudes of the individual terms in equations (3.68) and (3.80), it may be shown that $E[e_1(n)e_8(n)] > MPE_{opt}$ for

(a) $g_p > g_m$ (for all values of other process and model parameters)

and

(b) $g_p < g_m$, provided $K_p \ge K_m$ and $T_m \ge T_p$ (O'Dwyer (1996m)).

The conditions in (b) are sufficient, rather than necessary conditions.

However, if a gradient method is used to determine g_p, then an additional restriction that the MPE function must be unimodal with a minimum MPE value occurring at $g_m = g_p$, is imposed. The conditions for unimodality may be determined by induction. It is assumed that the excitation signal is a square wave signal of amplitude ± 1 and of period equal to 100 samples (i.e. that $g_p < 50$). An outline of the inductive proof (provided in full by O'Dwyer (1996m)) is as follows:

(a) $g_p > g_m$: It may be proved, using equations (3.68), (3.79) and (3.80), that the MPE function at $g_m = g_p - 1$ is greater than the MPE function at $g_m = g_p$ for all parameter values. Similarly, using equations (3.68) and (3.79), it may be proved that the MPE function at $g_m = g_p - n - 1$ is always greater than the MPE function at $g_m = g_p - n$, provided

$$\begin{aligned}
&0.04\{K_m(1 - e^{-T_s/T_m})\frac{T_s}{T_m} + K_p(1 - e^{-T_s/T_p})[2(1 - e^{-T_s/T_m}) - \frac{T_s}{T_m}]\} \\
&+0.08K_p(1 - e^{-T_s/T_m})(e^{-T_s/T_p} - e^{-T_s/T_m})(1 - e^{-g_m T_s/T_m}) \\
&-K_p(e^{-T_s/T_p} - e^{-T_s/T_m})\frac{T_s}{T_m}\left[(1 - e^{-T_s/T_p})e^{-(g_m-1)T_s/T_p} r_{uu}(n + 1 - g_m) - 0.04(1 - e^{-(g_m-1)T_s/T_p})\right] \\
&>0 \qquad (3.81)
\end{aligned}$$

This is a necessary condition.

(b) $g_p < g_m$: It may be proved, using equations (3.68), (3.79) and (3.80), that the MPE function at $g_m = g_p + 1$ is greater than the MPE function at $g_m = g_p$, provided that the following sufficient conditions are obeyed:

$$K_p \ge K_m \text{ and } T_m \ge T_p \qquad (3.82)$$

Similarly, it may be proved, using equations (3.68) and (3.79), that the MPE function at $g_m = g_p + n + 1$ is greater than the MPE function at $g_m = g_p + n$, provided that

$$-0.04\{K_m(1 - e^{-T_s/T_m})\frac{T_s}{T_m} - K_p(1 - e^{-T_s/T_p})[2(1 - e^{-T_s/T_m}) - \frac{T_s}{T_m}]\}$$

$$-0.04K_p(e^{-T_s/T_p}-e^{-T_s/T_m})\{\frac{nT_s}{T_m}[e^{-nT_s/T_p}(2-e^{-(g_m-n+1)T_s/T_p})-1]$$

$$+[2(1-e^{-T_s/T_m})-\frac{nT_s}{T_m}][e^{-nT_s/T_p}(2-e^{-(g_m-n)T_s/T_p})-1]-\frac{T_s}{T_m}[e^{-(n-2)T_s/T_p}(2-e^{-(g_m-n+1)T_s/T_m})-1]\}$$

$$+K_p(e^{-T_s/T_p}-e^{-T_s/T_m})\frac{T_s}{T_m}[e^{-(g_m-1)T_s/T_p}(1-e^{-T_s/T_p})r_{uu}(g_m-n-1)]>0 \qquad (3.83)$$

This is a necessary condition. $\square$

The behaviour of the MPE function, given by equation (3.68), versus model time delay index is shown for representative simulations in Figures 3.34 and 3.35. In Figure 3.34, $K_p = 2.0$, $K_m = 1.0$, $T_p = 0.7$ seconds and $T_m = 1.0$ seconds, so that the conditions in equations (3.81), (3.82) and (3.83) are fulfilled; in Figure 3.35, $K_p = 2.0$, $K_m = 3.0$, $T_p = 0.7$ seconds and $T_m = 0.5$ seconds, so that the condition in equation (3.81) is fulfilled, but the conditions in equations (3.82) and (3.83) are violated. The normalised MPE (equal to the MPE divided by $r_{uu}(0)$) is plotted versus model time delay index in both cases, with $r_{ww}(0)$ put to zero and $g_p = 30$; the conditions taken are identical to those used to calculate the normalised MPE curves in Figures 3.28 and 3.29. The excitation signal used in the determination of Figures 3.34 and 3.35 is a square wave signal of amplitude ± 1 and of period equal to 100 samples. The results are as expected from the theorem.

Figure 3.34: Normalised MPE vs. time delay index - square wave excitation

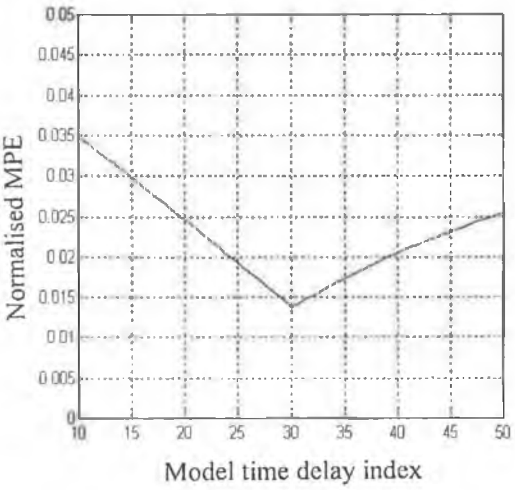
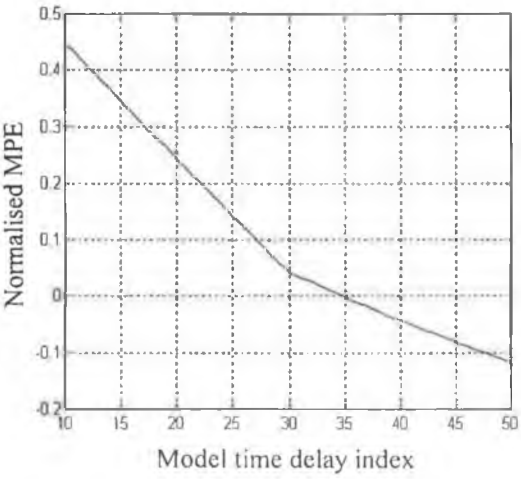


Figure 3.35: Normalised MPE vs. time delay index - square wave excitation



Considering equations (3.12), (3.13), (3.14), (3.53) and (3.54), a block diagram representation of the scheme to update both the model parameters and the model time delay index may be drawn; this block diagram is the same as Figure 3.30.

A representative simulation result corresponding to Theorem 3.7 is given in Figures 3.36a-3.36d, with the parameters and the process minus model output plotted against sample number. The simulation conditions for updating the time delay are identical to those in Section 3.5.2.1 (and thus, these simulation results may be compared with those in Figures 3.31a to 3.31d), except that the excitation signal is a square wave input of period equal to 100 samples and amplitude of ± 1. The normalised MPE corresponding to these conditions is given by Figure 3.34. In addition, the learning rate for the model time delay is put to 10 and filtering on the time delay update is employed.

Figure 3.36a: Gain estimate

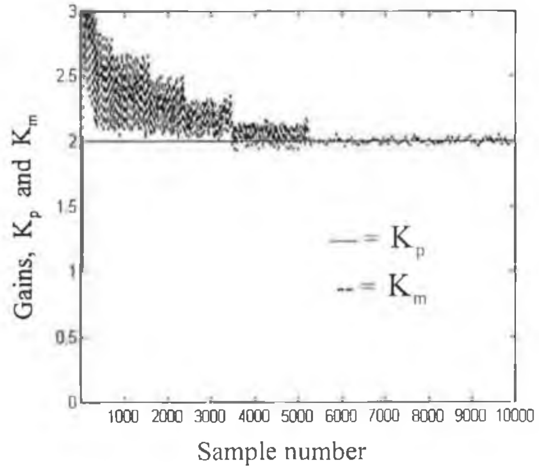


Figure 3.36b: Time constant estimate

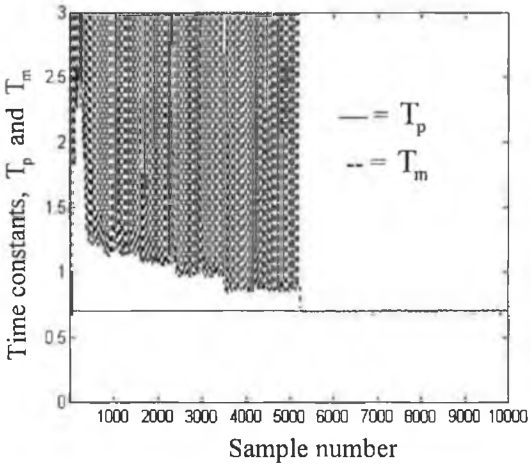


Figure 3.36c: Time delay index estimate

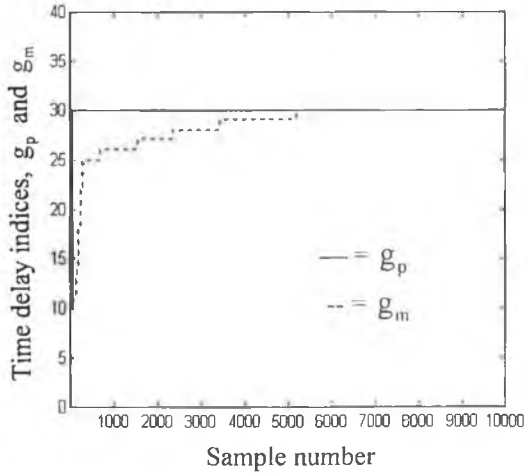
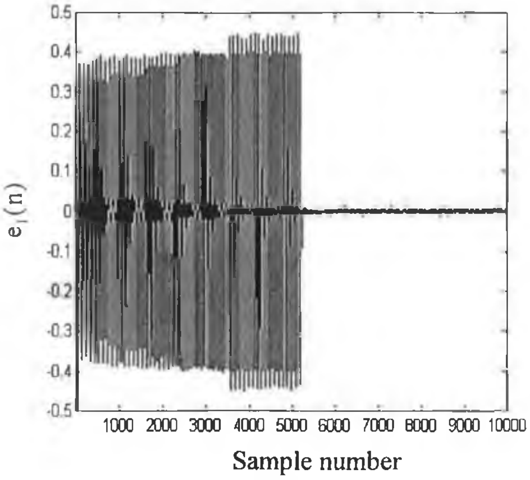


Figure 3.36d: $e_1(n)$



These results conform with Theorem 3.7.

3.5.2.5 The time delay as a real multiple of the sample period - square wave input

Theorem 3.7 has dealt with the estimation of time delays that are integer multiples of the sample period. For the estimation of time delays that are real multiples of the sample period, then the difference equation of a FOLPD process, $y_4(n)$, is given by equation (3.73). The corresponding model difference equation, $y_{m10}(n)$, is given by equation (3.74) (assuming the previous process output is used in its calculation). The model difference equation for calculating the partial derivative of the error with respect to the time delay variation, $y_{m11}(n)$, is given by equation (3.75) (assuming that the previous process output is used in its calculation). The error, $e_{10}(n) = y_4(n) - y_{m10}(n)$, is given by equation (3.76) and the error, $e_{11}(n) = y_4(n) - y_{m11}(n)$, is given by equation (3.77). The MPE performance surface, $E[e_{10}(n)e_{11}(n)]$, is given by equation (3.78).

The behaviour of the MPE function versus model time delay index, in representative simulation results, is given by Figures 3.37 and 3.38. For these simulations, $K_p = 2.0$, $K_m = 1.0$, $T_p = 0.7$ seconds, $T_m = 1.0$ seconds and $g_p = 5$, with the time delay taken in intervals of 0.01 times the sample period (i.e. the simulation conditions are identical to those in Section 3.5.2.2). The normalised MPE (equal to the MPE divided by $r_{uu}(0)$) is plotted versus time delay for $g_b = 0.0$ and $g_b = 0.5$. In Figure 3.37 the normalised MPE calculated from equation (3.68) is superimposed on the plots for comparison purposes. For all simulations, $r_{ww}(0)$ is put to zero. Clearly the MPE function is multimodal with respect to the time delay, when the time delay is a real multiple of the sample period (as in Section 3.5.2.2). The estimation of the real value of the time delay is therefore impossible using gradient methods.

Figure 3.37: Normalised MPE vs. time delay index - square wave excitation - $g_b = 0.0$

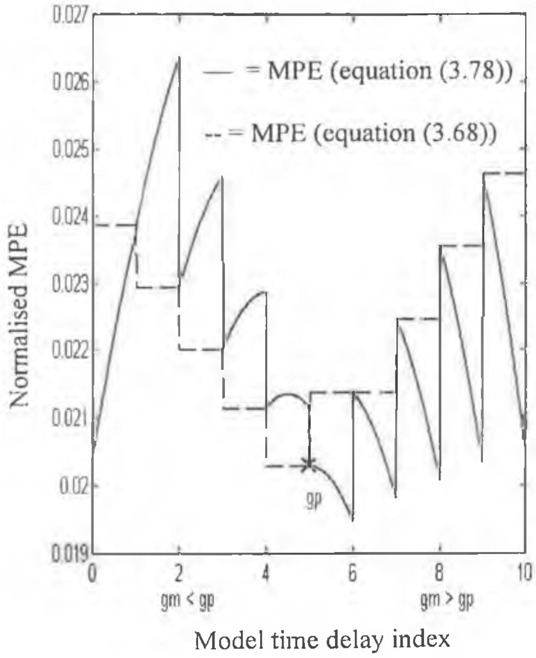
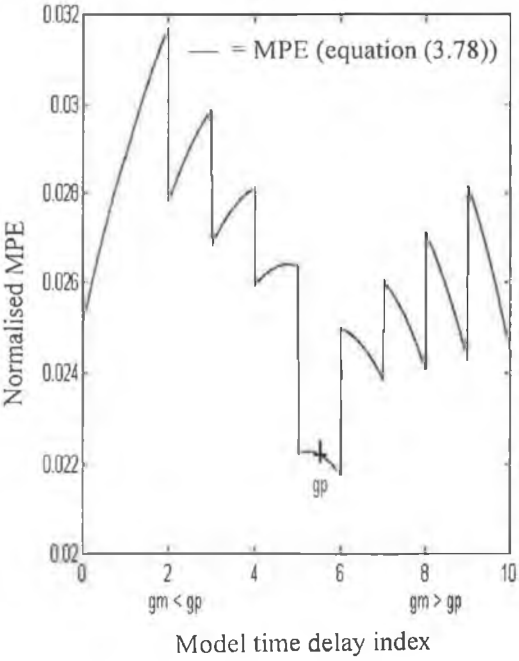


Figure 3.38: Normalised MPE vs. time delay index - square wave excitation - $g_b = 0.5$



3.5.2.6 Conclusions

Overall, the gradient method will allow the estimation of process parameters, for an appropriate square wave excitation signal, when a number of necessary and sufficient conditions on the process and model parameters are fulfilled. A non-gradient method that estimates the process time delay index by determining the minimum positive value of the MPE surface will also work under the same conditions. Unfortunately, it is not possible to estimate the process parameters if the process time delay is not an integer multiple of the sample period, using the gradient method; the indications are from Figures 3.37 and 3.38 that it may be possible to approximately estimate the process time delay if it is not an integer multiple of the sample period (at least in the simulations taken), using a non-gradient method that calculates the minimum positive value of the MPE cost function.

3.6 Conclusions

1. The most appropriate choice of rational polynomial to use to approximate the time delay variation of a process, modelled by an appropriate model, if the parameters and time delay of a FOLPD process are to be estimated using a gradient algorithm, is the first order Taylor's series approximation.

2. Seven theorems have been developed to analytically describe the conditions under which the model parameters may converge to the process parameters. These theorems are provided in the text of the chapter. Two of the theorems deal with the idealised case that the process time delay index is known *a priori*. The case when the time delay index is unknown *a priori* is also considered; unfortunately, the corresponding cost functions may be unimodal when $g_p > g_m$ only. This has meant that the time delay is often correctly estimated only because of appropriate filtering on the parameter. Some cases of unimodality do exist for all time delay index values (e.g. the simulation corresponding to Figure 3.34); however, various conditions indicated by equations (3.81) to (3.83) must be observed on the process and model parameters to achieve this result, which are impossible to evaluate prior to the implementation (as the process parameters are generally unknown). In addition, the inability of the relevant proposed methods to estimate time delays that are real multiples of the sample period (and, consequently, the corresponding process gain and time constants) is disappointing. Both of these features are difficult to reconcile with a practical application. The requirement that in some cases the excitation signal to the process should be of white noise form is another difficulty, as such a signal is not realisable in practice; however, other excitation signals may also be used, as described in the theorems. On a positive note, the fact that unimodality does exist on the cost function for some conditions, when the time delay is unknown *a priori*, provides some encouragement; it is particularly interesting that, comparing Figures 3.28 and 3.34, it is possible to achieve cost function unimodality with a square wave excitation signal and not with a white noise excitation signal, for one simulation condition. It would be interesting to filter the data before identification, as Ferretti *et al.* (1996) suggest that this may increase the range of time delay over which the cost function is unimodal, though the speed of convergence of any gradient algorithm used tends to be reduced. In addition, if the

process time delay index may be estimated accurately, an estimate of a process time delay that is a real multiple of the sample period could be determined by fitting an appropriate curve to a plot of the cost function (calculated, perhaps, in simulation) versus model time delay index.

3. An analytical approach, using the principle of induction, has been clearly outlined to allow the determination of the most appropriate excitation signal to use for a gradient based method for estimating the process time delay index and other parameters. The simulations taken are quite harsh tests of the algorithms, with the model parameters required to follow large step changes in the process parameters; as mentioned in the introduction, the algorithms may only be reasonably expected to work well for small differences between the process and model parameters. The main difficulty with the use of the gradient algorithms, as implemented, is the estimation of the time delay term. One avenue of future work that may be fruitful would be to estimate the time delay using an alternative approach, and estimate the non-delay parameters using the gradient approach. One such alternative approach would be to use a multiple model estimation method; it may also be possible to use an alternative gradient approach in the frequency domain, for example. Such a gradient approach is described in Chapter 4.

4. The work has concentrated on estimation, using gradient methods, when a FOLPD transfer function is used for both the process and the model. The investigation of the usefulness of the method if either the transfer function of the process or the model or both is not in FOLPD form could be carried out. Some preliminary work has already been done on the identification of the parameters of a second order model (with two poles and one zero) plus a time delay, when the time delay variation is approximated by a first order Taylor's series approximation and a first order Pade approximation. However, it is unlikely that a wider range of operating conditions for the gradient algorithms will be determined than those described for a FOLPD process and model in the theorems, particularly if the process and model are not of the same order. In addition, the development of appropriate theorems, corresponding to the theorems defined for the FOLPD case, will be more mathematically involved as the number of parameters to be estimated increases.

Chapter 4

Frequency domain methods of parameter and time delay estimation

4.1 Introduction

The estimation of the parameters (including the time delay) of a model in the frequency domain may be considered to be divided into two stages: firstly, the estimation of the process frequency response over an appropriate frequency range and secondly, the estimation of the parameters of the model from the frequency response. Both of these stages have been explored in detail in the published literature and are discussed in Chapter 2 of this thesis.

This chapter will first explore methods of estimating the frequency response of a process, both in open loop and in closed loop. The methods considered will be based on using the ratio of the Fourier transforms of output and input signals, and based on power spectral density techniques. Both of these methods promise to facilitate the estimation of the process frequency response in a relatively low number of computations (unlike the use of higher order spectral techniques, for instance, as discussed in Chapter 2). The estimation of the parameters of a model from the frequency response, by combining analytical methods and gradient based methods, will then be detailed. The analytical methods are based on direct calculation of the parameters from the frequency response, which provide initial parameter estimates. The gradient method then updates these initial estimates to more accurate model parameter estimates.

4.2 Process frequency response measurement

4.2.1 Introduction

The process frequency response for an open loop system is trivially calculated as the ratio of the Fourier transforms of the output and input signals to the process. One recursive technique for calculating the transforms is to use the Discrete Time Fourier Transform (DTFT), defined as follows:

$$F(\omega) = T_s \sum_{k=0}^{\infty} f(kT_s) e^{-j\omega k T_s} \tag{4.1}$$

with T_s = sample time. This transform has the advantage that a new term may be added as new data points become available. The DTFT could be modified by including tapering on the data window at the start and current evaluation points of the summation; this proposal would reduce spectral leakage. The inclusion of a non-rectangular data window would, however, increase the computational complexity of the calculation. An alternative recursive method for determining the transforms is to apply a numerical integration technique to the Fourier transform. An example of suitable techniques is the Adams-Moulton set, as discussed by Johnson and Reiss (1982). The first four elements of this set are as follows:

$$F_{k+1}(\omega) - F_k(\omega) = T_s x_{k+1} \tag{4.2}$$

$$F_{k+2}(\omega) - F_{k+1}(\omega) = 0.5T_s(x_{k+2} + x_{k+1}) \tag{4.3}$$

$$F_{k+3}(\omega) - F_{k+2}(\omega) = 0.083T_s(5x_{k+3} + 8x_{k+2} + x_{k+1}) \tag{4.4}$$

$$F_{k+4}(\omega) - F_{k+3}(\omega) = 0.042T_s(9x_{k+4} + 19x_{k+3} - 5x_{k+2} + x_{k+1}) \tag{4.5}$$

with

$$x_k = f(kT_s) e^{-j\omega k T_s} \tag{4.6}$$

Equations (4.2) and (4.3) are readily identified as the backward difference and trapezoidal rule (bilinear transform) respectively. If the identities commence from k = 0 and have zero initial conditions, the first four terms of the integrals in (4.2) and (4.3) become, respectively

$$T_s[x_0 + x_1 + x_2 + x_3] \tag{4.7}$$

and

$$T_s[0.5x_0 + x_1 + x_2 + 0.5x_3] \tag{4.8}$$

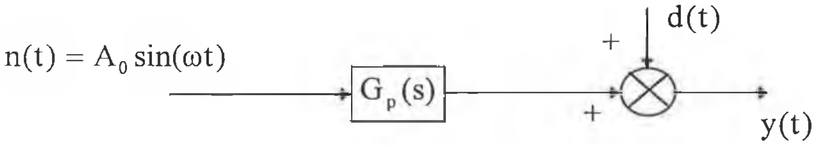
Note that equation (4.7) corresponds to a DTFT implementation. However, equation (4.8) corresponds to a DTFT implementation, with a data window which is tapered at each end. Higher order numerical integration techniques exaggerate this windowing effect.

4.2.2 Process frequency response identification in open loop

The identification of the open-loop frequency response of the process is aided by inputting to the process a sinusoidal excitation signal at the frequency at which the Fourier transform is being evaluated. The amplitude of this excitation signal should be commensurate with the amplitude of the measurement noise in the system.

Low pass filters on the magnitude and phase estimates may be used to reduce the effect of harmonic frequencies. These harmonic frequencies are seen to arise from the definition of the DTFT. The open loop system considered is represented as shown in Figure 4.1.

Figure 4.1: Open loop implementation



The process frequency response is estimated as follows:

$$G_p(j\omega) \approx \frac{Y(j\omega)}{N(j\omega)} \tag{4.9}$$

with $N(j\omega)$ and $Y(j\omega)$ being the Fourier transforms of n(t) and y(t), respectively. From Figure 4.1,

$$y(t) = A_0 \left| G_p(j\omega) \right| \sin(\omega t + \angle G_p(j\omega)) + d(t) \tag{4.10}$$

From Figure 4.1, equation (4.10) and assuming that the DTFT is used to implement $Y(j\omega)$ and $N(j\omega)$, then

$$N(j\omega) = T_s \sum_{k=0}^{\infty} A_0 [\sin(k\omega T_s)] e^{-jk\omega T_s} \tag{4.11}$$

i.e.

$$N(j\omega) = T_s A_0 \sum_{k=0}^{\infty} [\sin(k\omega T_s)][\cos(k\omega T_s - j\sin(k\omega T_s)] \tag{4.12}$$

i.e.

$$N(j\omega) = 0.5 T_s A_0 \sum_{k=0}^{\infty} [\sin(2k\omega T_s) - j(1 - \cos(2k\omega T_s))] \tag{4.13}$$

$Y(j\omega)$ is calculated, using the same procedure as that defined in equations (4.11) to (4.13), to be (O'Dwyer (1996m))

$$\begin{aligned} Y(j\omega) = {} & 0.5 T_s A_0 \left| G_p(j\omega) \right| \sum_{k=0}^{\infty} \left\{ \left[\sin \angle G_p(j\omega) + \sin(2k\omega T_s + \angle G_p(j\omega)) \right] + d(kT_s) \cos(k\omega T_s) \right\} \\ & - j 0.5 T_s A_0 \left| G_p(j\omega) \right| \sum_{k=0}^{\infty} \left\{ \left[\cos \angle G_p(j\omega) - \cos(2k\omega T_s + \angle G_p(j\omega)) \right] + d(kT_s) \sin(k\omega T_s) \right\} \end{aligned} \tag{4.14}$$

Therefore, from equations (4.13) and (4.14)

$$\frac{Y(j\omega)}{N(j\omega)} = \frac{\left| G_p(j\omega) \right| \sum_{k=0}^{\infty} \left\{ \left[\sin \angle G_p(j\omega) + \sin(2\omega k T_s + \angle G_p(j\omega)) \right] + d(kT_s) \cos(k\omega T_s) \right\}}{\sum_{k=0}^{\infty} \left\{ \sin(2\omega k T_s) + j(\cos(2\omega k T_s) - 1) \right\}}$$

$$-j\frac{|G_p(j\omega)|\sum_{k=0}^{\infty}\left\{\left[\cos\angle G_p(j\omega)-\cos(2\omega kT_s+\angle G_p(j\omega))\right]+d(kT_s)\sin(k\omega T_s)\right\}}{\sum_{k=0}^{\infty}\left\{\sin(2\omega kT_s)+j(\cos(2\omega kT_s)-1)\right\}} \tag{4.15}$$

The terms $\sin(2k\omega T_s)$ and $\cos(2k\omega T_s)$ in equation (4.15) show that harmonic frequencies at multiples of twice the DTFT frequency, ω, exists on the process frequency response estimate. If d(t) is assumed zero, then a difference equation for the phase of the process evaluated using the DTFT is directly calculated from equation (4.15) to be

$$\phi_k(j\omega)=\phi_{k-1}(j\omega)-\tan^{-1}\{\frac{\cos\angle G_p(j\omega)-\cos(2\omega kT_s+\angle G_p(j\omega))}{\sin\angle G_p(j\omega)+\sin(2\omega kT_s+\angle G_p(j\omega))}\}-\tan^{-1}\{\frac{-1+\cos 2\omega kT_s}{\sin 2\omega kT_s}\} \tag{4.16}$$

After convergence, $\phi_k(j\omega)=\phi_{k-1}(j\omega)=\angle G_p(j\omega)$ (on average), but the phase measurement continues to vary according to the latter two terms in equation (4.16), which involve the harmonic frequencies. However, when $\angle G_p(j\omega)=-\pi$, then from equation (4.16),

$$\phi_k(j\omega)=\phi_{k-1}(j\omega)-\tan^{-1}\{\frac{-1-\cos(2\omega kT_s-\pi)}{0+\sin(2\omega kT_s-\pi)}\}-\tan^{-1}\{\frac{-1+\cos 2\omega kT_s}{\sin 2\omega kT_s}\} \tag{4.17}$$

i.e. $\phi_k(j\omega)=\phi_{k-1}(j\omega)$ and the harmonic frequency terms are zero. In a similar manner, it may be demonstrated, using equation (4.15), that the magnitude variations due to the harmonic frequency terms are zero at $\angle G_p(j\omega)=-\pi$.

Low-pass filters on the magnitude and phase estimates may be used to reduce the effect of harmonic frequencies. The first order filters used are designed to have a cut-off frequency below 2ω. Alternatively, bandpass filters or filters with a variable cut-off frequency could be employed for improved performance.

The recursive schemes for calculating the frequency response need a form of data forgetting in a practical implementation. This is due to the constant adding of new terms as time progresses, which may cause the size of the DTFT to become very large. A consequence may be that the algorithm becomes insensitive to changes in the process dynamics or evaluation frequency, due to the magnitude difference between the new terms being added and the current transform size. An effective form of data forgetting is to weight the data values by progressively smaller amounts as they recede in time, by the introduction of an appropriate forgetting factor; for the first order DTFT with a rectangular data window, the formulation is as follows:

$$F_{k+1}(\omega) = \lambda F_k(\omega) + T_s x_{k+1} \tag{4.18}$$

with $0 < \lambda \leq 1$.

4.2.3 Frequency updating

The method described may be used to calculate the process frequency response at a number of different frequencies. If simple ultimate cycle methods for PID controller tuning are to be used, the calculation of the ultimate frequency (i.e. the frequency at a process phase of $-\pi$) is of interest; as detailed in Section 4.2.2, no harmonic frequencies exist on the process frequency response at this frequency. The method for calculating the ultimate frequency involves the adjustment of the evaluation frequency of the Fourier transform; it is proposed to extrapolate from previous phase and frequency values to determine the ultimate frequency. A number of data points may be used to fit a high order polynomial for the phase to the data, with the parameters of the polynomial being calculated using, for example, a least squares estimation strategy. The simplest algorithm of this type would be to fit a straight line to two data points, giving an updated estimate of the ultimate frequency as:

$$\omega_{i+1} = \omega_i - \delta \frac{\omega_i - \omega_{i-1}}{\phi_i - \phi_{i-1}} (\pi + \phi_i) \tag{4.19}$$

with $0 < \delta \leq 1$ and ω_{i-1}, ω_i and ω_{i+1} are frequencies corresponding to phases ϕ_{i-1}, ϕ_i and ϕ_{i+1}. δ may be considered to be an uncertainty factor that reflects the general non-linear nature of the phase response. If no *a priori* knowledge of the process is available, then a value of $\delta = 0.7$ gives a reasonable trade-off between the speed of convergence of the frequency towards the ultimate frequency and the phase response non-linearity.

The update regularity of the frequency estimate that is appropriate is related to the number of samples taken for the magnitude and phase measurements to settle, while retaining a reasonable rate of convergence of the frequency estimate to the ultimate frequency. A practical implementation allows updating every 100 samples for a trapezoidal integration technique, when the forgetting factor used is 0.97. The choice of forgetting factor is determined by a trade-off between convergence speed and noise immunity. A value as low as 0.8 may be used in a noise-free environment, giving rapid convergence and response to time varying systems, while a value closer to 0.99 is required to average out the effect of noise. The update regularity should therefore be chosen together with the forgetting factor.

Some simulation results showing the estimation of the magnitude and phase are detailed in Figures 4.2 to 4.4. The model used for the test is

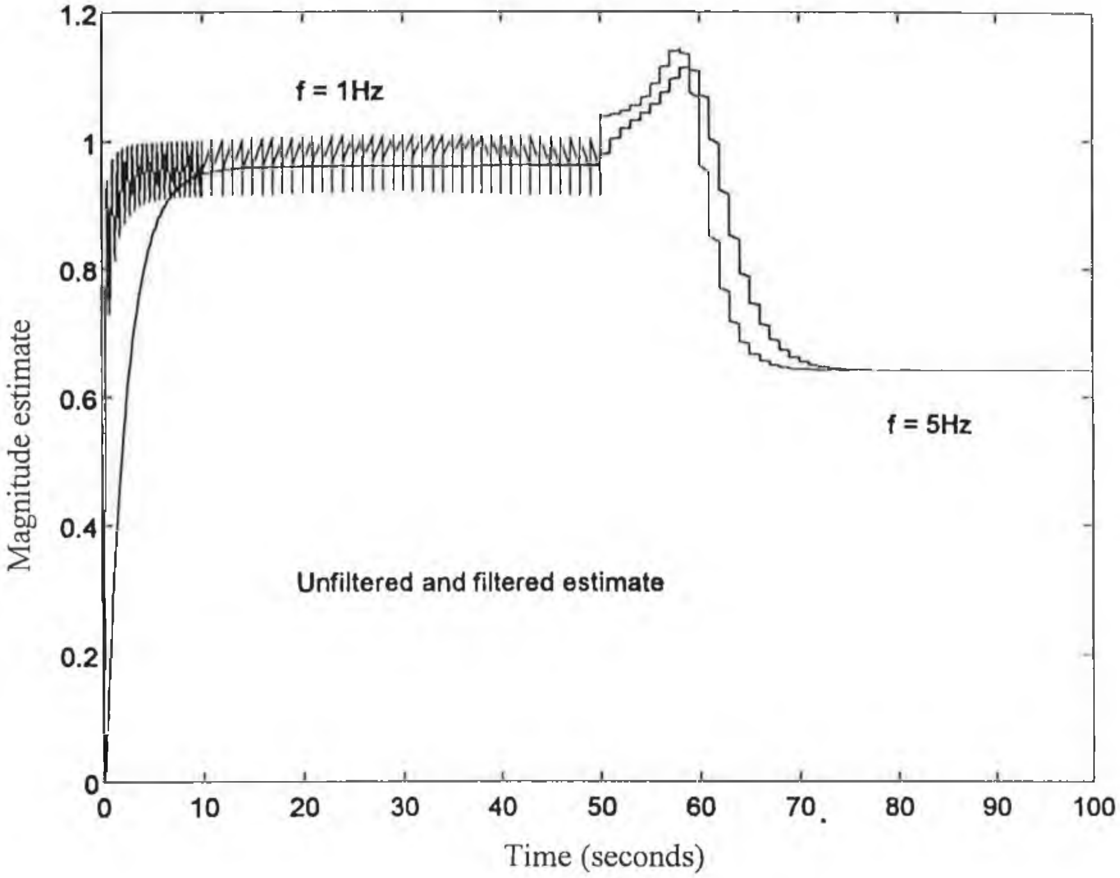
$$G_p(z) = \frac{0.805332z^{-1} + 0.100399z^{-2}}{1 - 0.096602z^{-1} + 0.002333z^{-2}} \tag{4.20}$$

This model is obtained by determining the discrete equivalent of $G_p(s) = e^{-0.1s}/(1+0.033s)^2$, taking the sample period to be 0.1 seconds. A trapezoidal integration method (equation (4.3)) is used for Fourier transform evaluation.

Figures 4.2 and 4.3a show the magnitude and phase estimates in open loop as the frequency is stepped from 1 Hz to 5 Hz. The low pass filter time constant in both cases is 10 seconds. The harmonic frequencies are evident in both cases when the estimates are unfiltered, though, as expected, the harmonic frequency terms tend to zero when the phase equals $-\pi$ radians (at 5 Hz). Figure 4.3b shows that the harmonic frequency fundamental (2 Hz) is twice the DTFT frequency (1 Hz). It has been calculated from equation (4.20) that the magnitude and phase of the process at 1 Hz are 0.96 and -0.75 radians, respectively, and the magnitude and phase at 5 Hz are 0.64 and -3.14 radians, respectively. These results compare well with those indicated in the figures.

Figure 4.4 shows the effect of varying the forgetting factor; reducing the forgetting factor allows faster convergence of the estimates at the expense of a larger variation in the unfiltered estimates.

Figure 4.2: Magnitude estimate - open loop - forgetting factor = 0.95



Figures 4.3a: Phase estimate - open loop - forgetting factor = 0.95

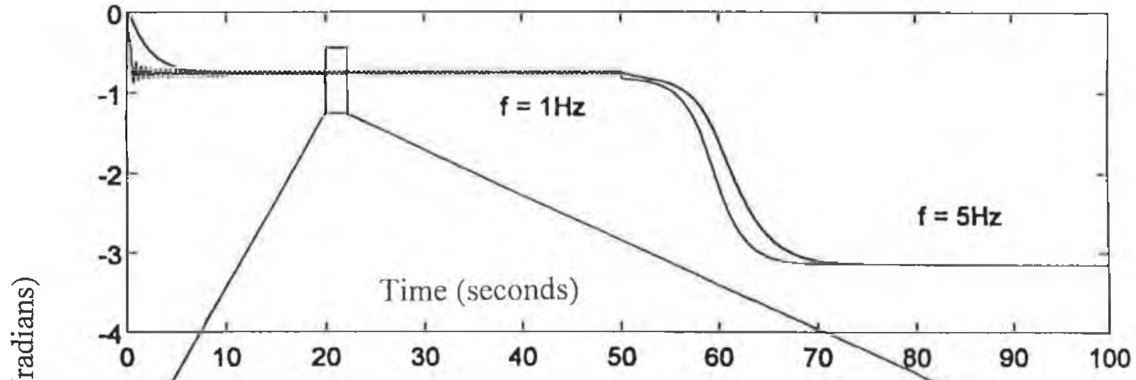


Figure 4.3b: Beat frequency (expanded)

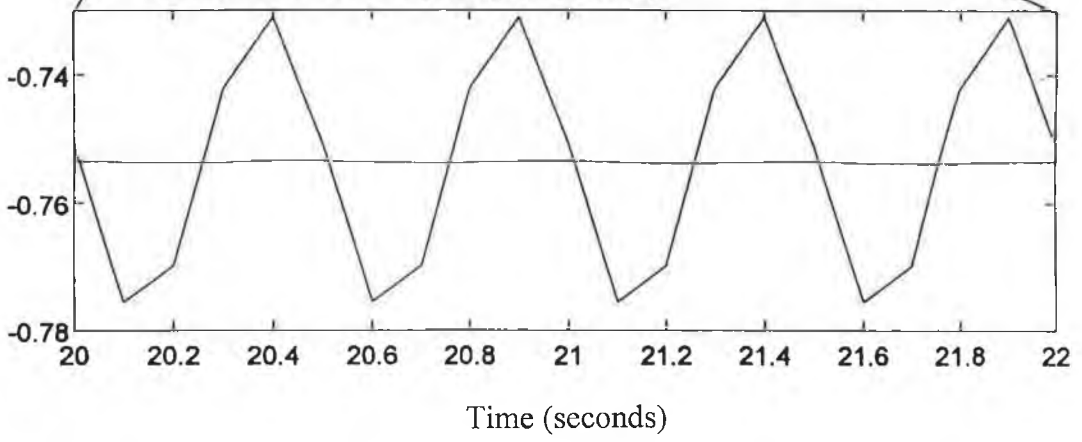
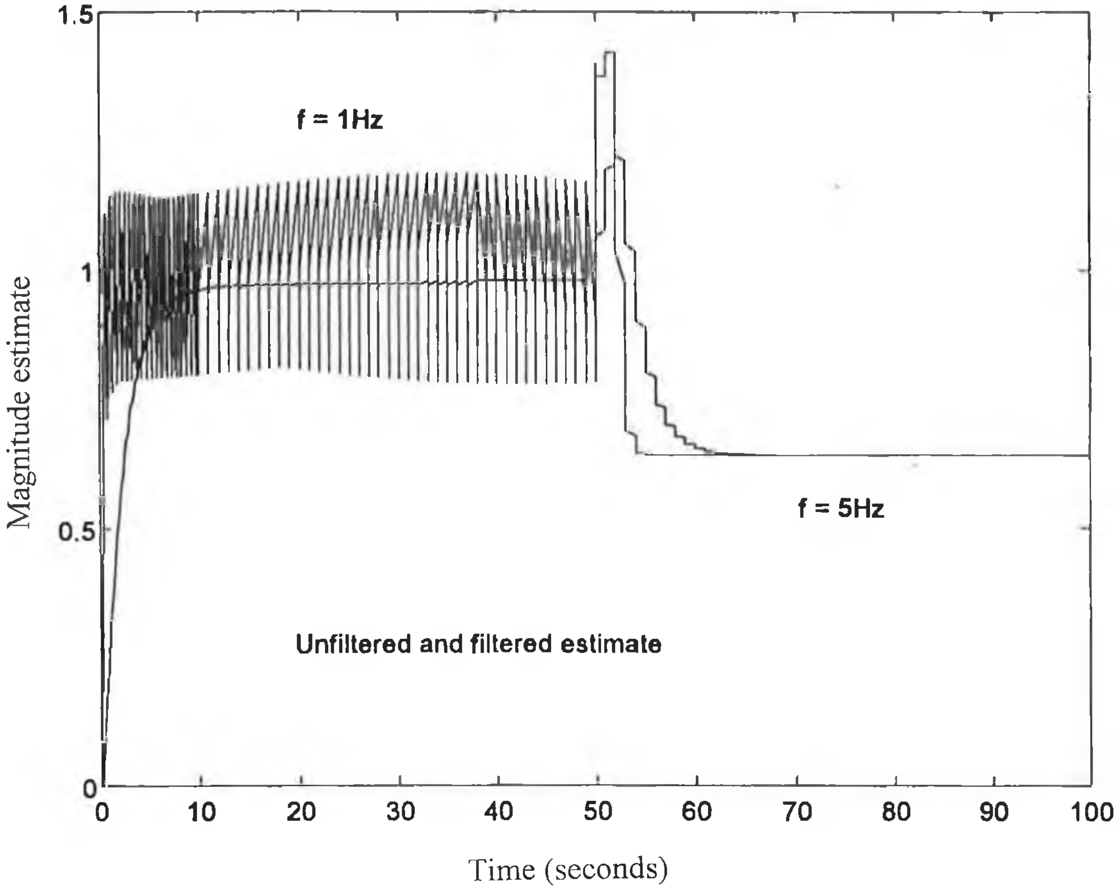


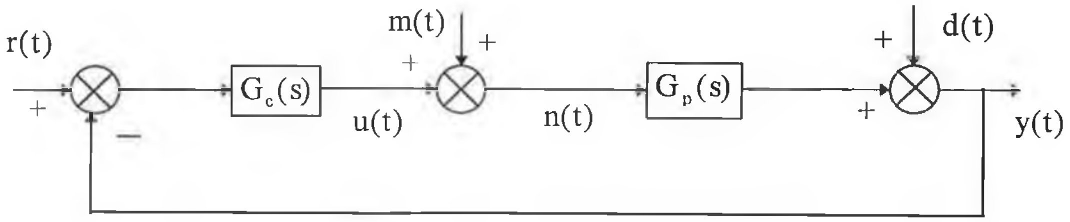
Figure 4.4: Magnitude estimate - open loop - forgetting factor = 0.8



4.2.4 Process frequency response identification in closed loop

The closed loop system considered is represented as shown in Figure 4.5.

Figure 4.5: Closed loop representation



If r(t), m(t) and d(t) are uncorrelated, then it may be deduced from Figure 4.5 that

$$n(s) = \frac{1}{1+G_c(s)G_p(s)}\left[G_c(s)(r(s) - d(s)) + m(s)\right] \tag{4.21}$$

It may also be concluded from Figure 4.5 that

$$y(s) = \frac{1}{1+G_c(s)G_p(s)}\left[d(s) + G_c(s)G_p(s)r(s) + G_p(s)m(s)\right] \tag{4.22}$$

Then, it may be calculated from equations (4.21) and (4.22) that

$$\frac{F[y(t)]}{F[n(t)]} = \frac{F[d(t)] + G_c(j\omega)G_p(j\omega)F[r(t)] + G_p(j\omega)F[m(t)]}{-G_c(j\omega)F[d(t)] + G_c(j\omega)F[r(t)] + F[m(t)]} \tag{4.23}$$

In the special case that F[d(t)] = 0, equation (4.23) reduces to

$$G_p(j\omega) = F[y(t)]/F[n(t)] \tag{4.24}$$

The Fourier transform terms may be approximated by using an appropriate integration method (equations (4.1) to (4.5)). In the development, the DTFT is used to implement the Fourier transform terms, and it is assumed that the effect of d(t) is neglected, by using appropriate filtering on n(t) and y(t). Then, it may be determined from Figure 4.5 and equation (4.24), using a procedure similar to that implemented in equations (4.10) to (4.15), that (O'Dwyer (1996m))

$$G_p(j\omega) \approx A_1\frac{a_4 + jb_4}{a_5 + jb_5} \tag{4.25}$$

with

$$a_4 = \sum_{k=0}^{\infty}\left\{\frac{A_2}{2}\left[\sin(2\omega kT_s + \phi_1 + \phi_2) + \sin(\phi_1 + \phi_2)\right] + r(kT_s)A_3\cos(\phi_1 + \phi_3 - \omega kT_s)\right\} \tag{4.26}$$

$$b_4 = \sum_{k=0}^{\infty}\left\{\frac{A_2}{2}\left[\cos(2\omega kT_s + \phi_1 + \phi_2) - \cos(\phi_1 + \phi_2)\right] + r(kT_s)A_3\sin(\phi_1 + \phi_3 - \omega kT_s)\right\} \tag{4.27}$$

$$a_5 = \sum_{k=0}^{\infty}\left\{\frac{A_2}{2}\left[\sin\left(2\omega kT_s + \phi_2\right) + \sin\left(\phi_2\right)\right] + r\left(kT_s\right)A_3\cos(\phi_3 - \omega kT_s)\right\} \tag{4.28}$$

and

$$b_5 = \sum_{k=0}^{\infty}\left\{\frac{A_2}{2}\left[\cos\left(2\omega kT_s + \phi_2\right) - \cos\left(\phi_2\right)\right] + r\left(kT_s\right)A_3\sin(\phi_3 - \omega kT_s)\right\} \tag{4.29}$$

and with

$$A_1 = \left|G_p(j\omega)\right|,\ \ \phi_1 = \angle[G_p(j\omega)] \tag{4.30}$$

$$A_2 = \left|P(j\omega)\right|,\ \ \phi_2 = \angle[P(j\omega)], \tag{4.31}$$

$$P(j\omega) = 1/[1 + G_p(j\omega)G_c(j\omega)] \tag{4.32}$$

and

$$A_3 = \left|Q(j\omega)\right|,\ \ \phi_3 = \angle[Q(j\omega)], \tag{4.33}$$

$$Q(j\omega) = G_c(j\omega)/(1 + G_p(j\omega)G_c(j\omega)) \tag{4.34}$$

The trigonometric terms in $2k\omega T_s$ in equations (4.26) to (4.29) show that harmonic frequencies at multiples of twice the DTFT frequency, ω, exist on the process frequency response estimate (equation (4.25)), as in the open loop case. A difference equation for the phase of the process, evaluated using the DTFT, is directly calculated from equations (4.25) to (4.29) to be (assuming $r(kT_s) = 0$)

$$\begin{aligned}\phi_k(j\omega) = \phi_{k-1}(j\omega) + \tan^{-1}\left\{\frac{-\cos(\phi_1 + \phi_2) + \cos(2\omega kT_s + \phi_1 + \phi_2)}{\sin(\phi_1 + \phi_2) + \sin(2\omega kT_s + \phi_1 + \phi_2)}\right\} \\ - \tan^{-1}\left\{\frac{-\cos(\phi_2) + \cos(2\omega kT_s + \phi_2)}{\sin(\phi_2) + \sin(2\omega kT_s + \phi_2)}\right\}\end{aligned} \tag{4.35}$$

After convergence, $\phi_k(j\omega) = \phi_{k-1}(j\omega) = \phi_1(= \angle G_p(j\omega))$ (on average), but the phase measurement continues to vary according to the latter two terms in equation (4.35), which involve the harmonic frequencies. However, when $\angle G_p(j\omega) = -\pi$, then from equation (4.35),

$$\phi_k(j\omega) = \phi_{k-1}(j\omega) + \tan^{-1}\{\frac{-\cos(\phi_2 - \pi) + \cos(2\omega kT_s + \phi_2 - \pi)}{\sin(\phi_2 - \pi) + \sin(2\omega kT_s + \phi_2 - \pi)}\} - \tan^{-1}\{\frac{-\cos(\phi_2) + \cos(2\omega kT_s + \phi_2)}{\sin(\phi_2) + \sin(2\omega kT_s + \phi_2)}\} \quad (4.36)$$

i.e. $\phi_k(j\omega) = \phi_{k-1}(j\omega)$ and the harmonic frequency terms are zero (as in the open loop case). In a similar manner, it may be demonstrated, using equations (4.25) to (4.29), that the magnitude variations due to the harmonic frequency terms are zero at $\angle G_p(j\omega) = -\pi$.

To aid identification in closed-loop, m(t) has been set up as an excitation signal at the appropriate (Fourier transform) frequency. This signal, while not having any adverse effects on the regulation properties of the system, would appear to be sufficient to allow consistent identification of the process frequency response in closed-loop.

A further practical addition of bandpass filters (on n(t) and y(t)) with moveable centre frequency may be included to concentrate calculations on the frequency range of interest. This helps to improve the disturbance and noise rejection properties of the adaptation algorithm. A Butterworth filter design is used with transfer function:

$$G_{bp}(z) = \frac{z^2 - 1}{z^2 - (1+\alpha)\beta z + \alpha} \quad (4.37)$$

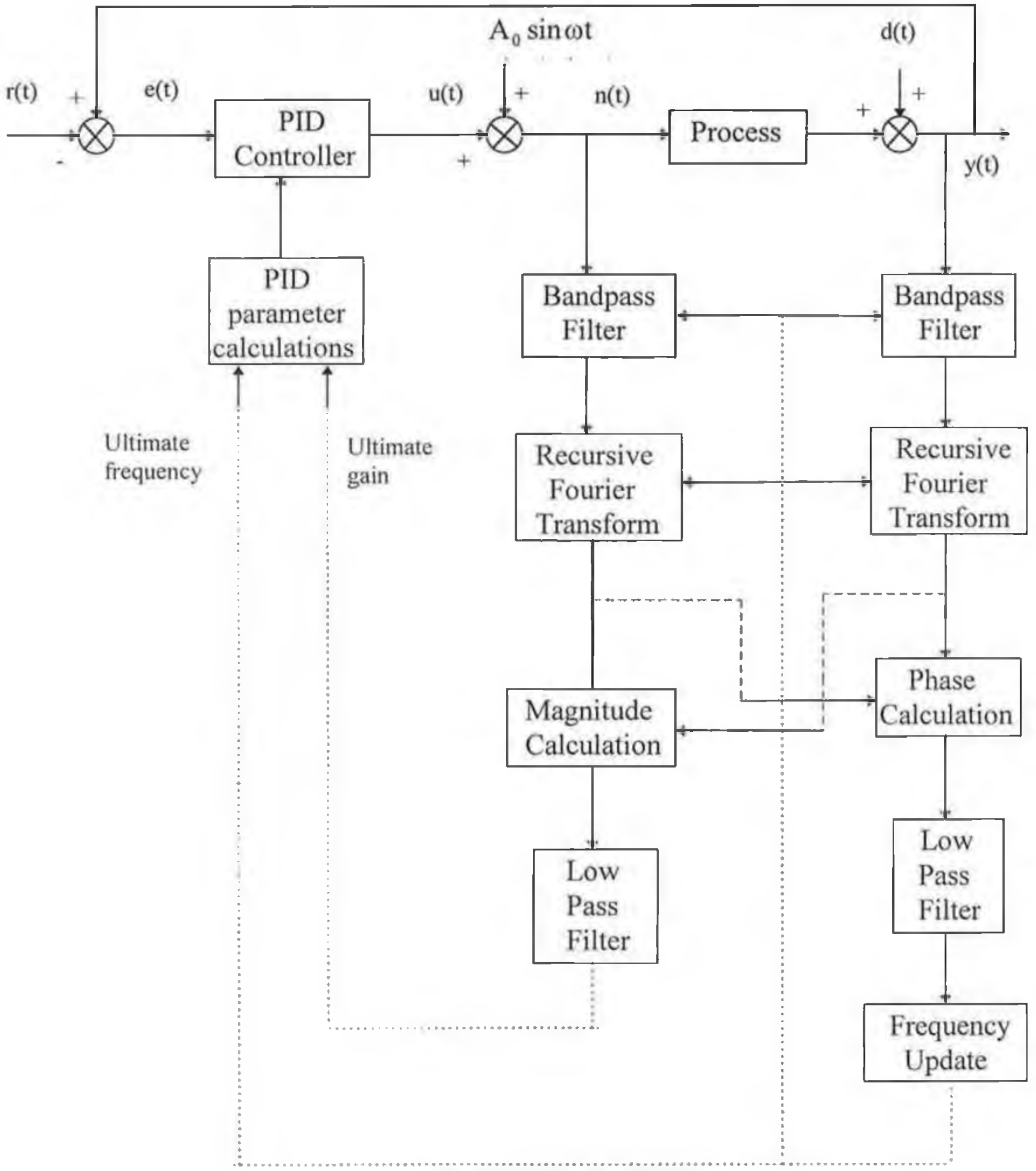
where

$$\beta = \cos(\omega_c T_s)/\cos(0.5\omega_{bw} T_s) \quad (4.38)$$

α is a parameter determined from the equivalent low-pass design and depends only on the filter bandwidth, ω_{bw}, and the sampling period, T_s; ω_c is the centre frequency of the bandpass filter (Lynn and Fuerst (1994)). Alternatively, a number of bandpass filters could be placed on the input and output of the process to determine the frequency response at a number of frequencies, corresponding to the centre frequencies of the bandpass filters.

The block diagram of the closed loop system implementation is shown in Figure 4.6; Ringwood and O'Dwyer (1994a), (1994b) use the ultimate gain and ultimate frequency to calculate the parameters of a PID compensator, as indicated in the figure.

Figure 4.6: Block diagram of the closed loop system implementation



A simulation result showing the estimation of the magnitude, phase and ultimate frequency is shown in Figure 4.7. The simulated model is

$$G_p(z) = \frac{0.11138z^{-1} + 0.09911z^{-2}}{1 - 1.684z^{-1} + 0.7047z^{-2}} z^{-2} \tag{4.39}$$

In the simulation: The sample period, T_s, equals 0.2 seconds

The forgetting factor, λ, equals 0.95

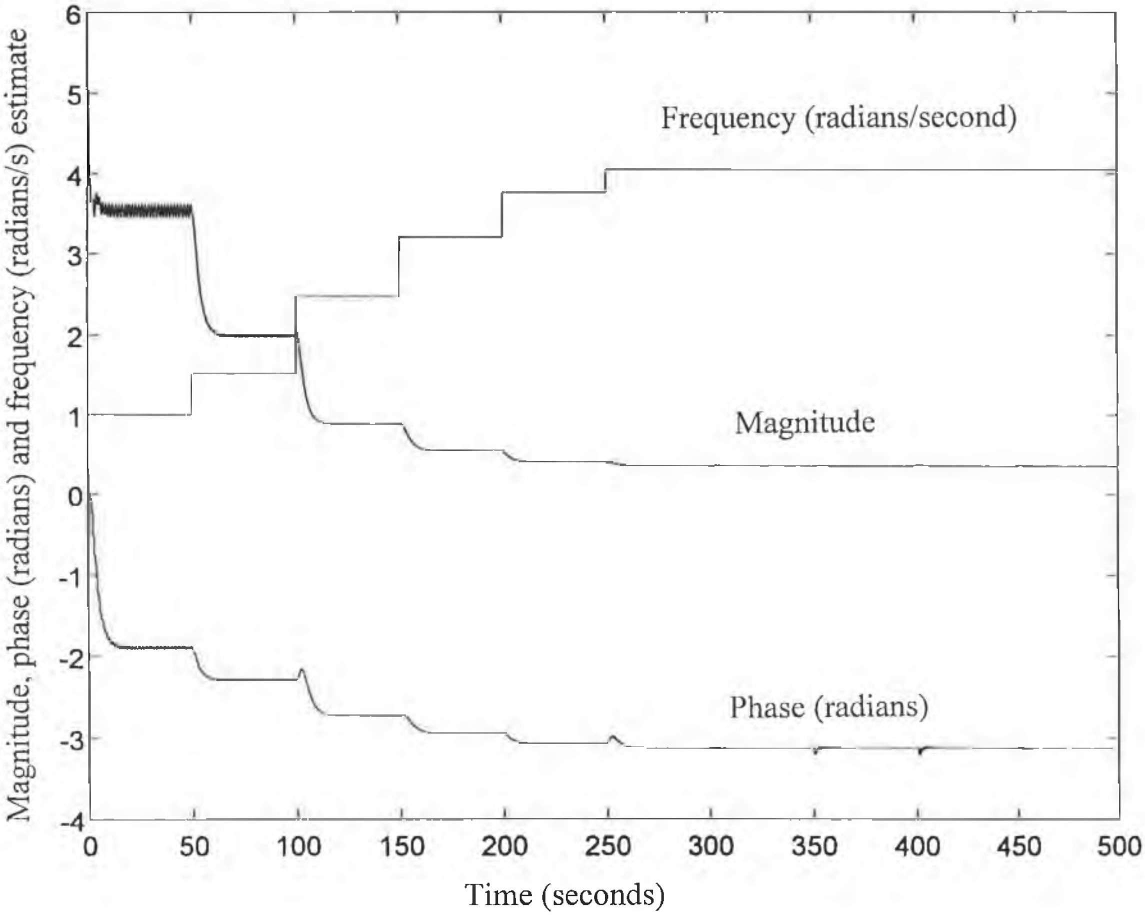
The low pass filter time constant equals 10 seconds

The estimates are updated every 500 samples

The bandpass filter bandwidth, ω_{bw}, equals 3 radians/second

The bandpass filter parameter, α, equals -0.7387.

Figure 4.7: Magnitude, phase and frequency convergence



The magnitude and phase values calculated at $\omega = 4$ radians/second (using equation (4.39)) are 0.34 and -3.15 radians, respectively; these results correspond quite well with the corresponding part of Figure 4.7.

4.2.5 Use of power spectral methods for identifying the process frequency response

Power spectral methods may be used as an alternative to Fourier transform methods to identify the frequency response of the process. The use of power spectral methods for the identification of the process frequency response in open loop has been discussed in Chapter 2; this section will consider the estimation of the process frequency response, using power spectral methods, in a closed loop environment.

The closed loop system considered is represented as shown in Figure 4.5 (Section 4.2.4). The power spectral density function and cross-power spectral density function are defined as follows:

$$S_n(j\omega) = \int_{-\infty}^{\infty} R_n(T)e^{-j\omega T}dT \tag{4.40}$$

and

$$S_{ny}(j\omega) = \int_{-\infty}^{\infty} R_{ny}(T)e^{-j\omega T}dT \tag{4.41}$$

with
$$R_n(T) = \lim_{T1\to\infty}\frac{1}{2T1}\int_{-T1}^{T1} n(t)n(t+T)dt \tag{4.42}$$

$$R_{ny}(T) = \lim_{T1\to\infty}\frac{1}{2T1}\int_{-T1}^{T1} n(t)y(t+T)dt \tag{4.43}$$

It may be deduced from Figure 4.5 that

$$u(s) = \frac{G_c(s)}{1 + G_c(s)G_p(s)}\left[r(s) - d(s) - G_p(s)m(s)\right] \tag{4.44}$$

It may subsequently be proved, by using the definitions in equations (4.40) to (4.43), and equations (4.22) and (4.44), and assuming r(t) = d(t) = 0, that

$$G_p(j\omega) = S_{uy}(j\omega)/S_u(j\omega) \tag{4.45}$$

Wellstead (1986) shows that, if r(t) = d(t) = 0,

$$G_p(j\omega) = S_{ny}(j\omega)/S_n(j\omega) \tag{4.46}$$

Wellstead (1986) also shows that, if r(t), m(t) and d(t) are uncorrelated, then

$$G_p(j\omega) = S_{my}(j\omega)/S_{mn}(j\omega) \tag{4.47}$$

or

$$G_p(j\omega) = S_{ry}(j\omega)/S_{m}(j\omega) \tag{4.48}$$

Approximations for the power spectral densities may be calculated by using the DTFT, for instance.

The estimate of the process frequency response, when the power spectral density is approximated by an appropriate DTFT, may be calculated by using equation (4.47), for example, together with the power spectral density and cross power spectral density formulae (equations (4.40) to (4.43)). Using these equations, and equations (4.21) and (4.22), and assuming that r(t), d(t) and m(t) are uncorrelated (with $m(t) = A_0 \sin(\omega t)$), it may be calculated that (O'Dwyer (1996m))

$$G_p(j\omega) \approx A_1 \frac{a_1 + a_2 + j(b_1 + b_2)}{a_3 + jb_3} \tag{4.49}$$

with

$$a_1 = \frac{A_0 A_2}{2} \sum_{k=0}^{\infty} \cos(\phi_1 + \phi_2) + \cos(2\omega k T_s + \phi_1 + \phi_2) \tag{4.50}$$

$$a_2 = \frac{A_3 c_k}{2} \sum_{k=0}^{\infty} \cos(\phi_1 + \phi_3) + \cos(2\omega k T_s + \phi_1 + \phi_3) \tag{4.51}$$

$$b_1 = \frac{A_0 A_2}{2} \sum_{k=0}^{\infty} \sin(\phi_1 + \phi_2) - \sin(2\omega k T_s + \phi_1 + \phi_2) \tag{4.52}$$

$$b_2 = -\frac{A_3 c_k}{2} \sum_{k=0}^{\infty} \sin(\phi_1 + \phi_3) + \sin(2\omega k T_s + \phi_1 + \phi_3) \tag{4.53}$$

$$a_3 = \frac{A_0 A_2}{2} \sum_{k=0}^{\infty} \cos(\phi_2) + \cos(2\omega k T_s + \phi_2) + \frac{A_3 c_k}{2} \sum_{k=0}^{\infty} \cos(\phi_3) + \cos(2\omega k T_s + \phi_3) \tag{4.54}$$

$$b_3 = \frac{A_0 A_2}{2} \sum_{k=0}^{\infty} \sin(\phi_2) - \sin(2\omega k T_s + \phi_2) - \frac{A_3 c_k}{2} \sum_{k=0}^{\infty} \sin(\phi_3) + \sin(2\omega k T_s + \phi_3)$$

(4.55)

and with

$$c_k = \frac{\lim\limits_{T\to\infty}\frac{1}{T}\int_{-T}^{T}[r(t+\tau)-d(t+\tau)]\sin\omega\tau \; d\tau}{\cos\omega\tau} \tag{4.56}$$

with $\omega = k\omega_s$, ω_s= sampling frequency. As in Section 4.2.4, the trigonometric terms in $2k\omega T_s$ in equations (4.50) to (4.55) show that harmonic frequencies at multiples of twice the DTFT frequency, ω, exist on the process frequency response estimate (equation (4.49)). In a similar manner to the procedure in Section 4.2.4, and using equations (4.49) to (4.56), it may be shown that, if the process frequency response is determined at the ultimate frequency, then the estimate taken at this frequency is an unbiased estimate.

4.2.6 Conclusions

A method has been defined to estimate the process frequency response, in open loop and in closed loop, using the ratio of Fourier transforms of the output and input signals to the process. The algorithm may be extended to include explicit time delay estimation, since the characteristic behaviour of a time delay in the frequency domain (i.e. a linear phase lag variation with frequency) may be resolved from the overall magnitude and phase measurements. Such an extension is not possible with parametric time-domain schemes. An alternative method to estimate the process frequency response, in open loop and in closed loop, using the ratio of power spectral density functions has also been outlined. The implementation of this method is recommended in future work.

Higher order spectral methods may also be used to estimate the process frequency response, as an alternative to the spectral analysis methods discussed. Such techniques are explored in detail in Chapter 2. However, it was concluded that the use of higher order spectral techniques in system identification seems suited to a restrictive range of open loop problems, in which noise signals on the input and output to the process cannot be effectively dealt with by pre-processing.

4.3 Model parameter estimation using frequency response data

4.3.1 Introduction

This section of the chapter discusses the estimation of the parameters (including the time delay) of SISO process models from an appropriate number of arbitrarily specified points on the process frequency response. The process time delay affects the phase response of the process, but not its magnitude response; Dos Santos and De Carvalho (1990) and Koganezawa (1991), for instance, use this feature to separately estimate the non-delay parameters and the time delay. Lilja (1988) estimates the parameters of a FOLPD process model by estimating the non-delay parameters through the minimisation of an appropriate cost function; the time delay is estimated separately by calculating the global minimum of a non-unimodal cost function using a modified Newton-Raphson algorithm. All of these approaches have the disadvantage of separately estimating the non-delay parameters and the time delay; this leads to biased estimation of the time delay or difficulty in achieving reliable convergence of the time delay estimate to its optimum value.

These difficulties motivate an investigation of the possibility of estimating the non-delay and time delay parameters together. A two stage approach, combining an analytical approach and a gradient approach, will be defined for the estimation of the parameters of an arbitrary order model plus time delay. The analytical methods are based on direct calculation of the parameters from the frequency response, using simultaneous equations which provide initial parameter estimates. A least squares approach using a gradient algorithm, updates the initial estimates to more accurate model parameter estimates (a least squares approach to the problem was originally suggested by Palmor and Blau (1994)). All of the parameters (including the time delay) are estimated together. This two stage approach will rely on the analytical estimates being sufficiently accurate so that unimodality of the cost function (equal to the sum of the squares of the sampled errors between the process and model frequency responses) with respect to the parameter estimates, exists from the analytical estimates to the gradient estimates. These methods will be developed for the estimation of the parameters of a general order model and will be applied to the estimation of the parameters of a FOLPD model and a SOSPD model, amongst others. Techniques will also be developed for the estimation of an appropriate model order, and some of the trade-offs experienced in choosing the model order will be explored.

A number of simulations were performed to demonstrate the operation of the two stage method. The following simulated process transfer functions were taken:

Case A:
$$G_p(s) = \frac{2e^{-0.5s}}{1+s} \tag{4.57}$$

Case B:
$$G_p(s) = \frac{2e^{-s}}{1+4.5s+4.5s^2} \tag{4.58}$$

Case C:
$$G_p(s) = \frac{2e^{-s}}{1+8.5s+22.5s^2+18s^3} \tag{4.59}$$

Case D:
$$G_p(s) = \frac{2e^{-s}}{1+18s+137s^2+567s^3+1403s^4+2103s^5+1846s^6+856s^7+158s^8} \tag{4.60}$$

Case E:
$$G_p(s) = \frac{2e^{-s}}{1+s+s^2} \tag{4.61}$$

Case F:
$$G_p(s) = \frac{2(1+2.25s)e^{-s}}{1+8.5s+22.5s^2+18s^3} \tag{4.62}$$

Case G:
$$G_p(s) = \frac{2(1-2.25s)e^{-s}}{1+8.5s+22.5s^2+18s^3} \tag{4.63}$$

These simulations cover a reasonable range of processes, including high-order, underdamped and non-minimum phase processes.

4.3.2 The estimation of the parameters of an arbitrary order model (with time delay)

The analytical estimation of the parameters of such a model is explored in Section 4.3.2.1; the estimation of the parameters of the model, using a gradient approach (and commencing from the model parameter values calculated from the analytical approach) is explored in Section 4.3.2.2.

4.3.2.1 Estimation using an analytical approach

The estimates of the parameters of an v^{th} order model plus time delay using an analytical approach are obtained by calculating the non-delay parameters from an appropriate number of simultaneous equations, using data points on the magnitude response; the time delay is then calculated from one data point on the phase response. The transfer function of the v^{th} order model plus time delay is defined as follows (with $v \geq u$)

$$G_m(s) = \frac{K_m\left(1 + b_{1m}s + b_{2m}s^2 + \ldots\ldots. + b_{um}s^u\right)e^{-s\tau_m}}{1 + a_{1m}s + a_{2m}s^2 + \ldots\ldots. + a_{vm}s^v} \tag{4.64}$$

Alternatively, the transfer function may be defined as

$$G_m(s) = \frac{\left(b'_{0m} + b'_{1m}s + b'_{2m}s^2 + \ldots\ldots. + b'_{um}s^u\right)e^{-s\tau_m}}{1 + a_{1m}s + a_{2m}s^2 + \ldots\ldots. + a_{vm}s^v} \tag{4.65}$$

with a parameter vector

$$x_1 = \left[a_{1m} \quad a_{2m} \ldots\ldots.. a_{vm} \quad b'_{0m} \quad b'_{1m} \quad b'_{2m} \ldots\ldots.. b'_{um} \quad \tau_m\right]^T, x_1 \in \Re^{u+v+2} \tag{4.66}$$

In the frequency domain, the model transfer function is

$$G_m(j\omega) = \frac{\left(b'_{0m} + j\omega b'_{1m} - \omega^2 b'_{2m} - j\omega^3 b'_{3m} + \omega^4 b'_{4m} + \ldots\ldots..\right)e^{-j\omega\tau_m}}{1 + j\omega a_{1m} - \omega^2 a_{2m} - j\omega^3 a_{3m} + \omega^4 a_{4m} + \ldots\ldots..} \tag{4.67}$$

Therefore, from equation (4.67), the numerator term of $G_m(j\omega)$ may be written as

$$N_m(j\omega) = \left\{ \sum_{q=0}^{\mathrm{int}\left[\frac{u}{2}\right]} b'_{2qm}(-1)^q \omega^{2q} + j \sum_{q=1}^{\mathrm{int}\left[\frac{u+1}{2}\right]} b'_{(2q-1)m}(-1)^{q-1} \omega^{2q-1} \right\} \tag{4.68}$$

and the denominator term of $G_m(j\omega)$ may be written as

$$D_m(j\omega) = \sum_{r=0}^{\mathrm{int}\left[\frac{v}{2}\right]} a_{2rm}(-1)^r \omega^{2r} + j \sum_{r=1}^{\mathrm{int}\left[\frac{v+1}{2}\right]} a_{(2r-1)m}(-1)^{r-1} \omega^{2r-1} \tag{4.69}$$

with $a_{0m} = 1$, $\mathrm{int}\left[\frac{v}{2}\right] =$ integer part of $\frac{v}{2}$, $\mathrm{int}\left[\frac{u}{2}\right] =$ integer part of $\frac{u}{2}$.

Therefore, from equations (4.68) and (4.69), the magnitudes of the numerator and denominator terms may be written as

$$|N_m(j\omega)| = \sqrt{\left[\sum_{q=0}^{\mathrm{int}\left[\frac{u}{2}\right]} (-1)^q b'_{2qm} \omega^{2q}\right]^2 + \left[\sum_{q=1}^{\mathrm{int}\left[\frac{u+1}{2}\right]} (-1)^{q-1} b'_{(2q-1)m} \omega^{2q-1}\right]^2} \tag{4.70}$$

and

$$|D_m(j\omega)| = \sqrt{\left[\sum_{r=0}^{\mathrm{int}\left[\frac{v}{2}\right]} (-1)^r a_{2rm} \omega^{2r}\right]^2 + \left[\sum_{r=1}^{\mathrm{int}\left[\frac{v+1}{2}\right]} (-1)^{r-1} a_{(2r-1)m} \omega^{2r-1}\right]^2} \tag{4.71}$$

Now, from equation (4.67), $|G_m(j\omega)|^2$ may be written as

$$|G_m(j\omega)|^2 = \frac{\left(d_{0m} + d_{1m}\omega^2 + d_{2m}\omega^2 + \ldots\ldots + d_{um}\omega^{2u}\right)}{1 + c_{1m}\omega^2 + c_{2m}\omega^4 + \ldots\ldots + c_{vm}\omega^{2v}} \tag{4.72}$$

with

$$d_{0m} = b_{0m}^{'\,2} = K_m^{\,2} \tag{4.73}$$

$$d_{1m} = b_{1m}^{'\,2} - 2b_{0m}^{'}b_{2m}^{'} \tag{4.74}$$

$$d_{2m} = b_{2m}^{'\,2} + 2b_{0m}^{'}b_{4m}^{'} - 2b_{1m}^{'}b_{3m}^{'} \tag{4.75}$$

……………………………………

$$d_{(u-1)m} = b_{(u-1)m}^{'\,2} + 2\sum_{q=1}^{\mathrm{int}\left[\frac{u-1}{2}\right]} b_{(2q-1)m}^{'}b_{(2u-2q-1)m}^{'} - 2\sum_{q=1}^{\mathrm{int}\left[\frac{u}{2}\right]} b_{(2q-2)m}^{'}b_{(2u-2q)m}^{'} \quad \text{, u even} \tag{4.76}$$

$$d_{um} = b_{um}^{'\,2} + 2\sum_{q=1}^{\mathrm{int}\left[\frac{u}{2}\right]} b_{(2q-2)m}^{'}b_{(2u-2q+2)m}^{'} - 2\sum_{q=1}^{\mathrm{int}\left[\frac{u}{2}\right]} b_{(2q-1)m}^{'}b_{(2u-2q+1)m}^{'} \quad \text{, u even} \tag{4.77}$$

$$c_{1m} = a_{1m}^2 - 2a_{0m}a_{2m} \tag{4.78}$$

$$c_{2m} = a_{2m}^2 + 2a_{0m}a_{4m} - 2a_{1m}a_{3m} \tag{4.79}$$

……………………………………

$$c_{(v-1)m} = a_{(v-1)m}^2 + 2\sum_{r=1}^{\mathrm{int}\left[\frac{v-1}{2}\right]} a_{(2r-1)m}a_{(2v-2r-1)m} - 2\sum_{r=1}^{\mathrm{int}\left[\frac{v}{2}\right]} a_{(2r-2)m}a_{(2v-2r)m} \quad \text{, v even} \tag{4.80}$$

$$c_{vm} = a_{vm}^2 + 2\sum_{r=1}^{\mathrm{int}\left[\frac{v}{2}\right]} a_{(2r-2)m}a_{(2v-2r+2)m} - 2\sum_{r=1}^{\mathrm{int}\left[\frac{v}{2}\right]} a_{(2r-1)m}a_{(2v-2r+1)m} \quad \text{, v even} \tag{4.81}$$

A minimum of u+v+1 data points on the magnitude response are required to estimate the parameters. If just u+v+1 data points are taken, the vector of magnitude response values squared is

$$F_1 = \left[\left|G_p(j\omega_1)\right|^2 \ldots\ldots \left|G_p(j\omega_{u+v+1})\right|^2 \right]^T, F \in \Re^{u+v+1} \tag{4.82}$$

with $\left|G_p(j\omega)\right|$ = process magnitude at frequency ω. Then, from equations (4.72) and (4.82),

$$\begin{bmatrix} \left|G_p(j\omega_1)\right|^2 \\ \left|G_p(j\omega_2)\right|^2 \\ \vdots \\ \left|G_p(j\omega_u)\right|^2 \\ \left|G_p(j\omega_{u+1})\right|^2 \\ \vdots \\ \left|G_p(j\omega_{u+v+1})\right|^2 \end{bmatrix} = \begin{bmatrix} 1 & \omega_1^2 & \cdot\cdot & \omega_1^{2u} & -\omega_1^2\left|G_p(j\omega_1)^2\right| & \cdot\cdot & -\omega_1^{2v}\left|G_p(j\omega_1)\right|^2 \\ 1 & \omega_2^2 & \cdot\cdot & \omega_2^{2u} & -\omega_2^2\left|G_p(j\omega_2)\right|^2 & \cdot\cdot & -\omega_2^{2v}\left|G_p(j\omega_2)\right|^2 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & \omega_u^2 & \cdot\cdot & \omega_u^{2u} & -\omega_u^2\left|G_p(j\omega_u)\right|^2 & \cdot\cdot & -\omega_u^{2v}\left|G_p(j\omega_u)\right|^2 \\ 1 & \omega_{u+1}^2 & \cdot\cdot & \omega_{u+1}^{2u} & -\omega_{u+1}^2\left|G_p(j\omega_{u+1})\right|^2 & \cdot\cdot & -\omega_{u+1}^{2v}\left|G_p(j\omega_{u+1})\right|^2 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & \omega_{u+v+1}^2 & \cdot\cdot & \omega_{u+v+1}^{2u} & -\omega_{u+v+1}^2\left|G(j\omega_{u+v+1})\right|^2 & \cdot\cdot & -\omega_{u+v+1}^{2v}\left|G_p(j\omega_{u+v+1})\right|^2 \end{bmatrix} \begin{bmatrix} d_{0m} \\ d_{1m} \\ \vdots \\ d_{um} \\ c_{1m} \\ \vdots \\ c_{vm} \end{bmatrix}$$

(4.83)

Thus

$$\begin{bmatrix} d_{0m} \\ d_{1m} \\ \vdots \\ d_{um} \\ c_{1m} \\ \vdots \\ c_{vm} \end{bmatrix} = \begin{bmatrix} 1 & \omega_1^2 & \cdot\cdot & \omega_1^{2u} & -\omega_1^2\left|G_p(j\omega_1)^2\right| & \cdot\cdot & -\omega_1^{2v}\left|G_p(j\omega_1)\right|^2 \\ 1 & \omega_2^2 & \cdot\cdot & \omega_2^{2u} & -\omega_2^2\left|G_p(j\omega_2)\right|^2 & \cdot\cdot & -\omega_2^{2v}\left|G_p(j\omega_2)\right|^2 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & \omega_u^2 & \cdot\cdot & \omega_u^{2u} & -\omega_u^2\left|G_p(j\omega_u)\right|^2 & \cdot\cdot & -\omega_u^{2v}\left|G_p(j\omega_u)\right|^2 \\ 1 & \omega_{u+1}^2 & \cdot\cdot & \omega_{u+1}^{2u} & -\omega_{u+1}^2\left|G_p(j\omega_{u+1})\right|^2 & \cdot\cdot & -\omega_{u+1}^{2v}\left|G_p(j\omega_{u+1})\right|^2 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & \omega_{u+v+1}^2 & \cdot\cdot & \omega_{u+v+1}^{2u} & -\omega_{u+v+1}^2\left|G(j\omega_{u+v+1})\right|^2 & \cdot\cdot & -\omega_{u+v+1}^{2v}\left|G_p(j\omega_{u+v+1})\right|^2 \end{bmatrix}^{-1} \begin{bmatrix} \left|G_p(j\omega_1)\right|^2 \\ \left|G_p(j\omega_2)\right|^2 \\ \vdots \\ \left|G_p(j\omega_u)\right|^2 \\ \left|G_p(j\omega_{u+1})\right|^2 \\ \vdots \\ \left|G_p(j\omega_{u+v+1})\right|^2 \end{bmatrix}$$

(4.84)

The non-delay parameters of the model may subsequently be calculated from equations (4.73) to (4.81). The time delay of the model may be calculated (using equations (4.65), (4.68) and (4.69)) to be

$$\tau_m = \frac{1}{\omega}\left\{-\phi_p(j\omega) + \tan^{-1}\left[\frac{\sum_{q=1}^{\mathrm{int}\left[\frac{u+1}{2}\right]}(-1)^{q-1} b'_{(2q-1)m}\omega^{2q-1}}{\sum_{q=0}^{\mathrm{int}\left[\frac{u}{2}\right]}(-1)^{q} b'_{2qm}\omega^{2q}}\right] - \tan^{-1}\left[\frac{\sum_{r=1}^{\mathrm{int}\left[\frac{v+1}{2}\right]}(-1)^{r-1} a_{(2r-1)m}\omega^{2r-1}}{\sum_{r=0}^{\mathrm{int}\left[\frac{v}{2}\right]}(-1)^{r} a_{2rm}\omega^{2r}}\right]\right\}$$

(4.85)

with $\phi_p(j\omega)$ = process phase at frequency ω. A less computationally intense alternative to the procedure defined in equations (4.64) to (4.85) is to estimate the parameters of a v^{th} order model, with no numerator parameters, and a repeated pole. This model is defined as follows:

$$G_{m1}(s) = \frac{K_{m1}e^{-s\tau_{m1}}}{(1+sT_{m1})^v} \tag{4.86}$$

K_{m1} and T_{m1} are calculated from two simultaneous equations, by equating the magnitudes of the model transfer function (calculated from equation (4.86)), at two frequency values ω_1 and ω_2, to two corresponding data points on the process magnitude response, $|G_p(j\omega_1)|$ and $|G_p(j\omega_2)|$ i.e.

$$|G_p(j\omega_1)| = K_{m1}/(1+\omega_1{}^2 T_{m1}{}^2)^{0.5v} \tag{4.87}$$

and

$$|G_p(j\omega_2)| = K_{m1}/(1+\omega_2{}^2 T_{m1}{}^2)^{0.5v} \tag{4.88}$$

From equations (4.87) and (4.88), it is calculated that

$$K_{m1} = \sqrt[0.5v]{\frac{(\omega_1{}^2 - \omega_2{}^2)}{(\omega_1{}^2/|G_p(j\omega_2)|^{2/v}) - (\omega_2{}^2/|G_p(j\omega_1)|^{2/v})}} \tag{4.89}$$

Using equations (4.87), (4.88) and (4.89), it is calculated that

$$T_{m1} = \frac{1}{\omega}\sqrt{\frac{K_{m1}{}^{2/v}}{|G_p(j\omega)|^{2/v}} - 1}, \quad \omega = \omega_1 \text{ or } \omega_2 \tag{4.90}$$

From equation (4.86), it is calculated that

$$\tau_{m1} = \frac{1}{\omega}[-\phi_p(j\omega) - v\tan^{-1}(\omega T_{m1})] \tag{4.91}$$

Equations (4.90) and (4.91) are also provided by Isermann *et al.* (1974) (Chapter 2).

The lower computational intensity of this procedure is traded off against poorer accuracy of the parameters estimated, when compared to the analytical procedure used to define a general model with no numerator parameters, using equations (4.78) to (4.81) and equations (4.84) and (4.85).

4.3.2.2 Estimation using a gradient approach

The transfer function of the v^{th} order model is defined by equation (4.64). The parameter vector is

$$x_2 = \left[K_m \quad a_{1m} \quad a_{2m} \ldots\ldots\ldots a_{vm} \quad b_{1m} \quad b_{2m} \ldots\ldots\ldots b_{um} \quad \tau_m\right]^T, x_2 \in \Re^{u+v+2} \tag{4.92}$$

In the frequency domain, the model transfer function is

$$G_m(j\omega) = \frac{K_m\left(1 + j\omega b_{1m} - \omega^2 b_{2m} - j\omega^3 b_{3m} + \omega^4 b_{4m} + \ldots\ldots\ldots\right)e^{-j\omega\tau_m}}{1 + j\omega a_{1m} - \omega^2 a_{2m} - j\omega^3 a_{3m} + \omega^4 a_{4m} + \ldots\ldots\ldots} \tag{4.93}$$

The numerator term of $G_m(j\omega)$, $N_m(j\omega)$, is given by equation (4.68) and the denominator term of $G_m(j\omega)$, $D_m(j\omega)$, is given by equation (4.69). The magnitudes of the numerator and denominator terms ($|N_m(j\omega)|$ and $|D_m(j\omega)|$, respectively) are given by equations (4.70) and (4.71), respectively. The phase contributions of the numerator and denominator terms are calculated from equation (4.93) to be

$$\phi_m^N(j\omega) = \tan^{-1}\frac{\sum\limits_{q=1}^{\text{int}\left[\frac{u+1}{2}\right]} b_{(2q-1)m}(-1)^{q-1}\omega^{2q-1}}{\sum\limits_{q=0}^{\text{int}\left[\frac{u}{2}\right]} b_{2qm}(-1)^{q}\omega^{2q}} \tag{4.94}$$

and

$$\phi_m^D(j\omega) = -\tan^{-1}\frac{\sum_{r=1}^{\mathrm{int}\left[\frac{v+1}{2}\right]} a_{(2r-1)m}(-1)^{r-1}\omega^{2r-1}}{\sum_{r=0}^{\mathrm{int}\left[\frac{v}{2}\right]} a_{2rm}(-1)^{r}\omega^{2r}} \tag{4.95}$$

If u+v+1 data points on the frequency response are taken to estimate the parameters, the vector of frequency response values is

$$F_2 = \left[|G_p(j\omega_1)|......|G_p(j\omega_{u+v+1})| \quad \phi_p(j\omega_1).......\phi_p(j\omega_{u+v+1})\right]^T, F_2 \in \Re^{2u+2v+2} \tag{4.96}$$

The error vector is formed as follows:

$$e = \left[e_1 \quad e_2.......e_{u+v+1} \quad e_{u+v+2} \quad e_{u+v+3}........e_{2u+2v+2}\right]^T \tag{4.97}$$

with

$$e_n = \frac{|N_m(j\omega)|}{|D_m(j\omega)|} - |G_p(j\omega)| \quad , \quad 1 \le n \le u+v+1 \tag{4.98}$$

and

$$e_n = \phi_m^N(j\omega_{n1}) + \phi_m^D(j\omega_{n1}) - \omega_{n1}\tau_m - \phi_p(j\omega_{n1}) \quad , \quad u+v+1 < n \le 2u+2v+2$$
$$\text{and } n1 = n-u-v-1 \tag{4.99}$$

The cost function, J, is formulated as

$$J = 0.5e^T P e \tag{4.100}$$

with

$$P = \mathrm{diag}\left[\frac{1}{|G_p(j\omega_1)|} \quad \frac{1}{|G_p(j\omega_2)|}......\frac{1}{|G_p(j\omega_{u+v+1})|} \quad \frac{1}{\omega_1} \quad \frac{1}{\omega_2}......\frac{1}{\omega_{u+v+1}}\right] \tag{4.101}$$

The normalising matrix, P, is used to increase the range of parameters over which unimodality of the cost function exists. The cost function, J (using equations (4.98) to

(4.101)) may be calculated to be

$$J = 0.5\sum_{n=1}^{u+v+1}\left[\left(\frac{|N_m(j\omega_n)|}{|D_m(j\omega_n)|} - |G_p(j\omega_n)|\right)^2 \frac{1}{|G_p(j\omega_n)|} + \frac{1}{\omega_n}(\phi_m^N(j\omega_n) + \phi_m^D(j\omega_n) - \omega_n\tau_m - \phi_p(j\omega_n))^2\right] \tag{4.102}$$

Then, the updated estimate of the parameters at sample (k+1) may be calculated from the estimates at sample k, using the gradient algorithm:

$$x_2(k+1) = x_2(k) - \mu\frac{\partial J}{\partial x_2(k)} \tag{4.103}$$

with μ = learning rate. The initial values of the parameter estimates are determined using either of the analytical techniques detailed in Section 4.3.2.1. If $|N_m(j\omega)|$ (equation (4.70)) is formulated as

$$|N_m(j\omega)| = K_m\sqrt{\left[\sum_{q=0}^{\operatorname{int}\left[\frac{u}{2}\right]}(-1)^q b_{2qm}\omega^{2q}\right]^2 + \left[\sum_{q=1}^{\operatorname{int}\left[\frac{u+1}{2}\right]}(-1)^{q-1} b_{(2q-1)m}\omega^{2q-1}\right]^2} \tag{4.104}$$

with $b_{0m} = 1$, then it is clear from equations (4.102) and (4.104) that the cost function is quadratic in the gain estimate, K_m (for all values of the numerator and denominator parameters and time delay estimates). It is also clear from equation (4.102) that the cost function is quadratic in the time delay estimate, τ_m (for all values of the gain, numerator and denominator parameter estimates). The cost function is not, however, quadratic in the estimates of the other numerator and denominator parameter values, as may be deduced from equations (4.71), (4.94), (4.95), (4.102) and (4.104). The cost function must be unimodal with respect to each of these parameter values (allowing the time delay estimate, gain estimate and other parameter estimates to vary), and must have its minimum value when the appropriate equivalent process parameter equals the model parameter, if convergence of the model parameters to the equivalent process parameters is to be guaranteed. An equivalent condition is that the first partial derivative of the cost function with respect to each of the parameter values may be equal to zero once only, or that the second partial derivative of the cost function with respect to each of the parameter values must always be greater than zero (with the first partial derivative of the cost function with respect to each of the parameter values being equal to zero at appropriate parameter values). Therefore, defining

$$N_{m1}(j\omega) = \left\{ \sum_{q=0}^{\operatorname{int}\left[\frac{u}{2}\right]} b_{2qm}(-1)^{q}\omega^{2q} \right\} \tag{4.105}$$

$$N_{m2}(j\omega) = \left\{ \sum_{q=1}^{\operatorname{int}\left[\frac{u+1}{2}\right]} b_{(2q-1)m}(-1)^{q-1}\omega^{2q-1} \right\} \tag{4.106}$$

$$D_{m1}(j\omega) = \sum_{r=0}^{\operatorname{int}\left[\frac{v}{2}\right]} a_{2rm}(-1)^{r}\omega^{2r} \tag{4.107}$$

and

$$D_{m2}(j\omega) = \sum_{r=1}^{\operatorname{int}\left[\frac{v+1}{2}\right]} a_{(2r-1)m}(-1)^{r-1}\omega^{2r-1} \tag{4.108}$$

the first partial derivatives with respect to the parameters may be calculated, using equations (4.71), (4.94), (4.95), (4.102) and equations (4.104) to (4.108). If h is even, the unimodality conditions are

$$\frac{\partial J}{\partial a_{(h-1)m}} = -0.5\sum_{n=1}^{u+v+1} \frac{D_{m2}(j\omega_n)(-1)^{\frac{h}{2}-1}\omega_n{}^{h-1}\left|N_m(j\omega_n)\right|}{\left|D_m(j\omega_n)\right|^3}\left[\frac{\left|N_m(j\omega_n)\right|}{\left|D_m(j\omega_n)\right|} - \left|G_p(j\omega_n)\right|\right]$$
$$-0.5\sum_{n=1}^{u+v+1} \frac{(-1)^{\frac{h}{2}-1}\omega_n{}^{h-1}D_{m1}(j\omega_n)}{\left|D_m(j\omega_n)\right|^2}\left[\phi_m^N(j\omega_n) + \phi_m^D(j\omega_n) - \omega_n\tau_m - \phi_p(j\omega_n)\right] = 0 \tag{4.109}$$

once only and

$$\frac{\partial J}{\partial a_{hm}} = -0.5\sum_{n=1}^{u+v+1}\frac{D_{m1}(j\omega_n)(-1)^{\frac{h}{2}}\omega_n{}^{h}\left|N_m(j\omega_n)\right|}{\left|D_m(j\omega_n)\right|^3}\left[\frac{\left|N_m(j\omega_n)\right|}{\left|D_m(j\omega_n)\right|}-\left|G_p(j\omega_n)\right|\right]$$

$$-0.5\sum_{n=1}^{u+v+1}\frac{(-1)^{\frac{h}{2}+1}\omega_n{}^{h}D_{m2}(j\omega_n)}{\left|D_m(j\omega_n)\right|^2}\left[\phi_m^N(j\omega_n)+\phi_m^D(j\omega_n)-\omega_n\tau_m-\phi_p(j\omega_n)\right]=0$$

(4.110)

once only. Also, if i is even,

$$\frac{\partial J}{\partial b_{(i-1)m}} = 0.5\sum_{n=1}^{u+v+1}\frac{K_m{}^2N_{m2}(j\omega_n)(-1)^{\frac{i}{2}-1}\omega_n{}^{i-1}}{\left|D_m(j\omega_n)\right|\left|N_m(j\omega_n)\right|}\left[\frac{\left|N_m(j\omega_n)\right|}{\left|D_m(j\omega_n)\right|}-\left|G_p(j\omega_n)\right|\right]$$

$$+0.5\sum_{n=1}^{u+v+1}\frac{(-1)^{\frac{i}{2}-1}\omega_n{}^{i-1}N_{m1}(j\omega_n)K_m{}^2}{\left|N_m(j\omega_n)\right|^2}\left[\phi_m^N(j\omega_n)+\phi_m^D(j\omega_n)-\omega_n\tau_m-\phi_p(j\omega_n)\right]=0$$

(4.111)

once only and

$$\frac{\partial J}{\partial b_{im}} = 0.5\sum_{n=1}^{u+v+1}\frac{K_m{}^2N_{m1}(j\omega_n)(-1)^{\frac{i}{2}}\omega_n{}^{i}}{\left|D_m(j\omega_n)\right|\left|N_m(j\omega_n)\right|}\left[\frac{\left|N_m(j\omega_n)\right|}{\left|D_m(j\omega_n)\right|}-\left|G_p(j\omega_n)\right|\right]$$

$$+0.5\sum_{n=1}^{u+v+1}\frac{K_m{}^2(-1)^{\frac{i}{2}+1}\omega_n{}^{i}N_{m2}(j\omega_n)}{\left|N_m(j\omega_n)\right|^2}\left[\phi_m^N(j\omega_n)+\phi_m^D(j\omega_n)-\omega_n\tau_m-\phi_p(j\omega_n)\right]=0$$

(4.112)

once only. The second partial derivatives with respect to the parameters may be calculated from equations (4.109) to (4.112). If h is even, the unimodality conditions are

$$\frac{\partial^2 J}{\partial a^2{}_{(h-1)m}} = \sum_{n=1}^{u+v+1}\frac{\left[2D_{m2}{}^2(j\omega_n)-D_{m1}{}^2(j\omega_n)\right]\omega_n{}^{2h-2}\left|N_m(j\omega_n)\right|}{2\left|D_m(j\omega_n)\right|^5}\left[\frac{\left|N_m(j\omega_n)\right|}{\left|D_m(j\omega_n)\right|}-\left|G_p(j\omega_n)\right|\right]^2$$

$$+\sum_{n=1}^{u+v+1}\frac{2\omega_n{}^{2h-2}D_{m1}(j\omega_n)D_{m2}(j\omega_n)}{2\left|D_m(j\omega_n)\right|^4}\left[\phi_m^N(j\omega_n)+\phi_m^D(j\omega_n)-\omega_n\tau_m-\phi_p(j\omega_n)\right]$$

$$+\sum_{n=1}^{u+v+1}\left\{\frac{D_{m2}{}^2(j\omega_n)\omega_n{}^{2h-2}\left|N_m(j\omega_n)\right|^2}{2\left|D_m(j\omega_n)\right|^6}+\frac{D_{m1}{}^2(j\omega_n)\omega_n{}^{2h-2}}{2\left|D_m(j\omega_n)\right|^4}\right\}>0 \qquad (4.113)$$

with $\partial J/\partial a_{(h-1)m} = 0$ at an appropriate value of $a_{(h-1)m}$ and

$$\frac{\partial^2 J}{\partial a^2{}_{hm}} = \sum_{n=1}^{u+v+1} \frac{\left[2D_{m1}{}^2(j\omega_n) - D_{m2}{}^2(j\omega_n)\right]\omega_n{}^{2h}\left|N_m(j\omega_n)\right|}{2\left|D_m(j\omega_n)\right|^5}\left[\frac{\left|N_m(j\omega_n)\right|}{\left|D_m(j\omega_n)\right|} - \left|G_p(j\omega_n)\right|\right]^2$$
$$- \sum_{n=1}^{u+v+1} \frac{2\omega_n{}^{2h}D_{m1}(j\omega_n)D_{m2}(j\omega_n)}{2\left|D_m(j\omega_n)\right|^4}\left[\phi_m^N(j\omega_n) + \phi_m^D(j\omega_n) - \omega_n\tau_m - \phi_p(j\omega_n)\right]$$
$$+ \sum_{n=1}^{u+v+1} \left\{\frac{D_{m1}{}^2(j\omega_n)\omega_n{}^{2h}\left|N_m(j\omega_n)\right|^2}{2\left|D_m(j\omega_n)\right|^6} + \frac{D_{m2}{}^2(j\omega_n)\omega_n{}^{2h}}{2\left|D_m(j\omega_n)\right|^4}\right\} > 0 \qquad (4.114)$$

with $\partial J/\partial a_{hm} = 0$ at an appropriate value of a_{hm}. Also, if i is even,

$$\frac{\partial^2 J}{\partial b^2{}_{(i-1)m}} = \sum_{n=1}^{u+v+1} \frac{K_m{}^4\left[N_{m1}{}^2(j\omega_n)\right]\omega_n{}^{2i-2}}{2\left|N_m(j\omega_n)\right|^3\left|D_m(j\omega_n)\right|}\left[\frac{\left|N_m(j\omega_n)\right|}{\left|D_m(j\omega_n)\right|} - \left|G_p(j\omega_n)\right|\right]$$
$$- \sum_{n=1}^{u+v+1} \frac{2\omega_n{}^{2i-2}N_{m1}(j\omega_n)N_{m2}(j\omega_n)}{2\left|N_m(j\omega_n)\right|^4}\left[\phi_m^N(j\omega_n) + \phi_m^D(j\omega_n) - \omega_n\tau_m - \phi_p(j\omega_n)\right]$$
$$+ \sum_{n=1}^{u+v+1} \left\{\frac{K_m{}^4N_{m1}{}^2(j\omega_n)\omega_n{}^{2i-2}}{2\left|D_m(j\omega_n)\right|^2\left|N_m(j\omega_n)\right|^2} + \frac{N_{m1}{}^2(j\omega_n)\omega_n{}^{2i-2}}{2\left|N_m(j\omega_n)\right|^4}\right\} > 0 \qquad (4.115)$$

with $\partial J/\partial b_{(i-1)m} = 0$ at an appropriate value of $b_{(i-1)m}$ and

$$\frac{\partial^2 J}{\partial b^2{}_{im}} = \sum_{n=1}^{u+v+1} \frac{K_m{}^4\left[N_{m2}{}^2(j\omega_n)\right]\omega_n{}^{2i}}{2\left|N_m(j\omega_n)\right|^3\left|D_m(j\omega_n)\right|}\left[\frac{\left|N_m(j\omega_n)\right|}{\left|D_m(j\omega_n)\right|} - \left|G_p(j\omega_n)\right|\right]$$
$$+ \sum_{n=1}^{u+v+1} \frac{2\omega_n{}^{2i}N_{m1}(j\omega_n)N_{m2}(j\omega_n)}{2\left|N_m(j\omega_n)\right|^4}\left[\phi_m^N(j\omega_n) + \phi_m^D(j\omega_n) - \omega_n\tau_m - \phi_p(j\omega_n)\right]$$
$$+ \sum_{n=1}^{u+v+1} \left\{\frac{K_m{}^4N_{m1}{}^2(j\omega_n)\omega_n{}^{2i}}{2\left|D_m(j\omega_n)\right|^2\left|N_m(j\omega_n)\right|^2} + \frac{N_{m2}{}^2(j\omega_n)\omega_n{}^{2i}}{2\left|N_m(j\omega_n)\right|^4}\right\} > 0 \qquad (4.116)$$

with $\partial J/\partial b_{im} = 0$ at an appropriate value of b_{im}.

Unfortunately, it was not possible to prove the conditions represented by equations (4.109) to (4.116) analytically, either in the general case or for any particular model structure.

4.3.3 Case studies

Four case studies are presented to demonstrate the wide applicability of the analytical and gradient methods presented in Sections 4.3.2.1 and 4.3.2.2, respectively. These case studies are

(1) The estimation of the parameters of a FOLPD model

(2) The estimation of the parameters of a SOSPD model (with no zero)

(3) The estimation of the parameters of a third order process plus delay model (with no zero) and

(4) The estimation of the parameters of a SOSPD model (with a zero)

4.3.3.1 FOLPD model parameter estimation

The transfer function of the model is defined as

$$G_m(s) = \frac{K_m e^{-s\tau_m}}{1 + sT_m} \tag{4.117}$$

From equations (4.73), (4.78), (4.84) and (4.85) respectively, the parameters K_m, T_m and τ_m of the FOLPD model may be calculated analytically, as follows:

$$K_m = \frac{|G_p(j\omega_1)||G_p(j\omega_2)|\sqrt{\omega_2^2 - \omega_1^2}}{\sqrt{|G_p(j\omega_2)|^2 \omega_2^2 - |G_p(j\omega_1)|^2 \omega_1^2}} \tag{4.118}$$

$$T_m = \frac{1}{\omega}\sqrt{\frac{K_m{}^2}{|G_p(j\omega)|^2} - 1}, \quad \omega = \omega_1 \text{ or } \omega_2 \tag{4.119}$$

$$\tau_m = \frac{1}{\omega}\left[-\phi_p(j\omega) - \tan^{-1}(\omega T_m)\right], \quad \omega = \omega_1 \text{ or } \omega_2 \tag{4.120}$$

(equations (4.89), (4.90) and (4.91) could also be applied to determine equations (4.118) to (4.120)). The sensitivity of the parameters K_m, T_m and τ_m with respect to the magnitude values recorded have been determined, by partially differentiating equations (4.118), (4.119) and (4.120) with respect to the magnitude values, to be:

$$\frac{\partial K_m}{\partial |G_p(j\omega_2)|} = \frac{-|G_p(j\omega_1)|^3 \omega_1^2 \sqrt{\omega_2^2 - \omega_1^2}}{\left[|G_p(j\omega_2)|^2 \omega_2^2 - |G_p(j\omega_1)|^2 \omega_1^2\right]^{1.5}} \tag{4.121}$$

$$\frac{\partial K_m}{\partial |G_p(j\omega_1)|} = \frac{|G_p(j\omega_2)|^3 \omega_2^2 \sqrt{\omega_2^2 - \omega_1^2}}{\left[|G_p(j\omega_2)|^2 \omega_2^2 - |G_p(j\omega_1)|^2 \omega_1^2\right]^{1.5}} \tag{4.122}$$

$$\frac{\partial T_m}{\partial |G_p(j\omega)|} = \frac{K_m}{\omega |G_p(j\omega)|^3 \sqrt{\left(\frac{K_m}{|G_p(j\omega)|}\right)^2 - 1}} \left[|G_p(j\omega)| \frac{\partial K_m}{\partial |G_p(j\omega)|} - K_m \right] \tag{4.123}$$

$$\frac{\partial \tau_m}{\partial |G_p(j\omega)|} = -\frac{1}{1 + \omega^2 T_m^2} \frac{\partial T_m}{\partial |G_p(j\omega)|} \tag{4.124}$$

with $|G_p(j\omega_1)|$ and $|G_p(j\omega_2)|$ being process magnitude values corresponding to ω_1 and ω_2; $\omega = \omega_1$ or ω_2.

It is evident from equations (4.121) and (4.122) that the sensitivity of the gain estimates to magnitude values recorded is reduced if K_m is calculated from magnitudes recorded far apart in frequency; numerical evaluations suggest that the magnitudes should be at least a decade apart in frequency. These numerical evaluations also suggest that:

(a) T_m should be calculated at frequencies when $0.25K_m < |G_p(j\omega)| < 0.75K_m$; it is interesting in this context that Sundaresan and Krishnaswamy (1978) state that, for good robustness, the time constant should be calculated when $|G_p(j\omega)| \approx 0.5K_m$.

(b) τ_m should be calculated at frequencies when $|G_p(j\omega)| < 0.5K_m$.

Ten values of the process frequency response, spaced equally between phase lags of 0^0 to 270^0 were recorded; equations (4.118) to (4.120) are used to calculate the FOLPD model parameters, for each of the processes indicated in Cases A to G (equations (4.57) to (4.63)). Average values of the parameters are calculated over a number of points of the frequency response (to improve the robustness of the estimates); the frequency response data values used in equations (4.118) to (4.120) were chosen to conform with the rules of thumb defined that reduce the sensitivity of the estimates to the magnitude values recorded. The model parameters were estimated under both noise free conditions and when either $+10\%$ or -10% was added to the magnitude and phase values of the process frequency response (labelled the $\pm 10\%$ noise condition). The estimates of the model parameters determined are summarised in Table 4.1.

Table 4.1: FOLPD model parameter estimates, calculated using the analytical method

	No Noise			±10% Noise		
	K_m	T_m	τ_m	K_m	T_m	τ_m
Case A	2.00	1.00	0.50	2.22	1.25	0.50
Case B	2.75	7.31	1.69	2.79	6.53	1.76
Case C	2.83	12.92	3.37	3.35	15.83	2.68
Case D	2.06	10.98	11.63	2.30	13.35	11.57
Case E	2.23	1.33	1.71	2.57	1.43	1.77
Case F	3.40	14.54	1.32	2.98	11.78	1.51
Case G	2.06	5.90	5.66	2.30	7.42	5.77

The large values of the sensitivity functions calculated in typical numerical evaluations means that overall, the analytical approach will provide what may be best regarded as the initial estimates of the parameters; poor estimates of the gain (which equals 2.00 in all processes taken) is seen in many of the results in Table 4.1, for instance, particularly under the $\pm 10\%$ noise condition.

The use of the gradient algorithm (equation (4.103) requires that the cost function be unimodal with respect to the time constant. The range of values of the parameters to ensure equation (4.113) is true (with $a_{(h-1)m} = T_m$) may be determined in simulation for the seven processes in Cases A to G under consideration (as it was not possible to calculate this range of values analytically). Equation (4.113) reduces (for this application) to

$$\frac{\partial^2 J}{\partial T_m{}^2} = 0.5\sum_{n=1}^{2}\{\frac{\omega_n{}^4 K_m{}^2 T_m{}^2}{|G_p(j\omega_n)|(1+\omega_n{}^2 T_m{}^2)^3} + \frac{\omega_n}{(1+\omega_n{}^2 T_m{}^2)^2}\}$$

$$+0.5\sum_{n=1}^{2}\{\frac{K_m \omega_n{}^2(2\omega_n{}^2 T_m{}^2 - 1)}{|G_p(j\omega_n)|(1+\omega_n{}^2 T_m{}^2)^{2.5}}\left[\frac{K_m}{\sqrt{1+\omega_n{}^2 T_m{}^2}} - |G_p(j\omega_n)|\right]\}$$

$$-0.5\sum_{n=1}^{2}\{\frac{2\omega_n{}^2 T_m}{(1+\omega_n{}^2 T_m{}^2)^2}\left[\tan^{-1}(\omega_n T_m) + \omega_n \tau_m + \phi_p(j\omega_n)\right]\} > 0 \qquad (4.125)$$

Numerical evaluation of equation (4.125) revealed that the initial estimates of the parameter values to be used in the gradient approach should be as defined, in either Option 1 or Option 2, to facilitate unimodality of the cost function with respect to the time constant variation.

Option 1: $K_m(\text{initial} - \text{gradient}) = 1.5K_m(\text{analytical})$ (4.126)

$$0.25T_m(\text{optimum}) \le T_m(\text{analytical}) \le 3.3T_m(\text{optimum}) \qquad (4.127)$$

$$\tau_m(\text{initial} - \text{gradient}) = 0.5\tau_m(\text{analytical}) \qquad (4.128)$$

with $K_m(\text{initial} - \text{gradient})$ and $\tau_m(\text{initial} - \text{gradient})$ being the initial values of the gain and time delay values, respectively, to be used with the gradient method. T_m (optimum) is the least squares value of the time constant calculated using the gradient method. $K_m(\text{analytical})$, $T_m(\text{analytical})$ and $\tau_m(\text{analytical})$ are the values of gain, time constant and time delay, respectively, calculated using the analytical approach.

Option 2: $0.83K_m(\text{optimum}) \le K_m(\text{analytical}) \le 1.17K_m(\text{optimum})$ (4.129)

$$0.25T_m(\text{optimum}) \le T_m(\text{analytical}) \le 1.25T_m(\text{optimum}) \qquad (4.130)$$

$$0 \leq \tau_m(\text{analytical}) \leq 1.1\tau_m(\text{optimum}) \tag{4.131}$$

with K_m (optimum) and τ_m (optimum) being the least squares values of the gain and time delay calculated using the gradient method.

It is easier for the conditions in Option 1 to be fulfilled in practice, as there is a tendency for $T_m(\text{analytical})$ to be greater than $1.25T_m(\text{optimum})$ (at least in the simulations taken). The specifications in Option 1 and Option 2 are worst case specifications i.e. it is possible that the parameter estimates may converge to their optimum values, using the gradient method, when the relevant parameter estimates calculated using the analytical approach fall outside the parameter ranges supplied.

The analytical estimates are first calculated (using equations (4.118), (4.119) and (4.120)); then the initial values of the gain and time delay for the gradient estimates of the parameters are put equal to 1.5 and 0.5 times the analytical gain and time delay estimates, respectively (assuming Option 1 is taken). This strategy increases the probability of convergence to the optimum values of the parameter estimates using the gradient method, though it does not guarantee such convergence. Simulation results show convergence to fixed FOLPD parameter estimates using the gradient method for all of the processes under discussion, in both the noise free and $\pm 10\%$ noise condition considered, in considerably less than 1000 iterations. These simulation results are summarised in Table 4.2; there is broad agreement between the parameters estimated in the noise free and $\pm 10\%$ noise condition.

Table 4.2: FOLPD model parameter estimates, calculated using the gradient method

	No Noise			±10% Noise		
	K_m	T_m	τ_m	K_m	T_m	τ_m
Case A	2.00	1.00	0.50	2.02	1.02	0.51
Case B	1.96	4.65	1.73	1.93	4.55	1.77
Case C	1.84	7.67	3.47	1.78	7.46	3.55
Case D	1.75	8.00	12.01	1.66	7.35	12.40
Case E	2.32	1.15	1.66	2.32	1.16	1.68
Case F	1.96	6.71	1.84	1.95	6.69	1.87
Case G	2.19	6.73	5.62	2.17	6.48	5.74

One representative simulation result is provided in Figures 4.8 to 4.13, for the process in Case C i.e. $G_p = 2e^{-s}/(1 + 8.5s + 22.5s^2 + 18s^3)$; the results are obtained under the ±10% noise condition.

Figure 4.8 shows that, for the simulation taken, a wide range of initial parameter values is possible (O = points where $\partial^2 J/\partial {T_m}^2 < 0$ and • = points where $\partial^2 J/\partial {T_m}^2 > 0$). The model parameter values calculated analytically in this case were $K_m = 3.35$, $T_m = 15.83$ seconds and $\tau_m = 2.68$ seconds (Table 4.1). Therefore, the initial estimates for the parameters when using the gradient algorithm are $K_m = 5.03$, $T_m = 15.83$ seconds and $\tau_m = 1.34$ seconds (Option 1); this estimate is marked as + on Figure 4.8.

Figures 4.9, 4.10 and 4.11 show the convergence of these parameter values to the optimum values within 500 samples, using the gradient method. The optimum value of $T_m = 7.46$ seconds means $T_m(\text{analytical}) = 2.12 T_m(\text{optimum})$, conforming with the guideline suggested in Option 1. Figures 4.12 and 4.13 show the step response and frequency response of the process and model together (using Program CC). The fitting of the process to the model in both domains is inaccurate (except at phase lags around 180^0), due primarily to an inaccurate estimate of the gain of the process. However, the apparent time delay of the process appears to be estimated well (Figure 4.12). Other simulation results show a similar deviation in the fitting between the process and the model, except when the process is itself of FOLPD structure (as may be deduced from the results provided for Case A in Table 4.2). It is possible, by restricting the range of phase values over which the process is identified, to yield a closer fitting between the process and the model in the frequency domain (over the corresponding frequency range) than that found in the simulation taken. Of course, the acceptability of the fitting of the model to the process in any particular frequency range depends on the use to which the model is applied.

Figure 4.8: Examination of $\partial^2 J/\partial T_m^2 < 0$ - $K_m = 5.03$

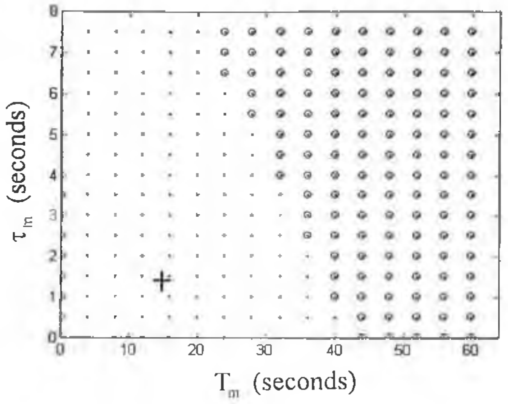


Figure 4.9: $K_m - \mu = 0.5$

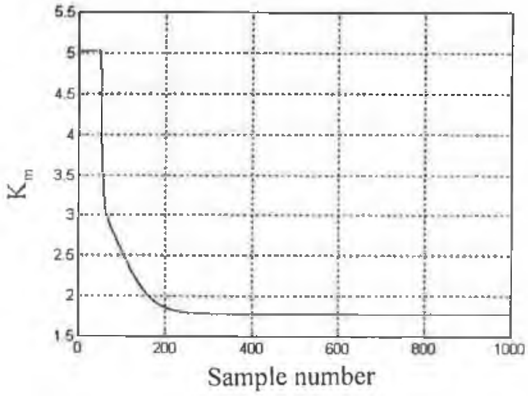


Figure 4.10: $T_m - \mu = 0.5$

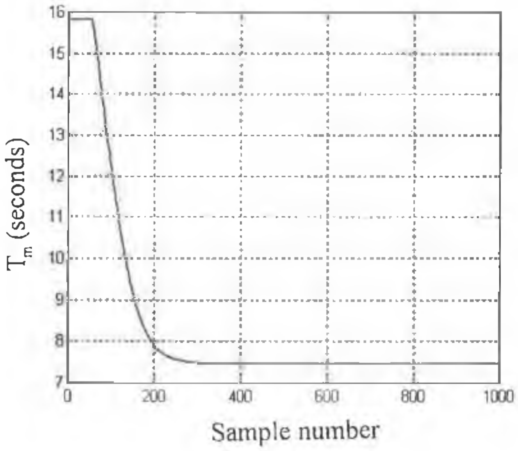


Figure 4.11: $\tau_m - \mu = 0.5$

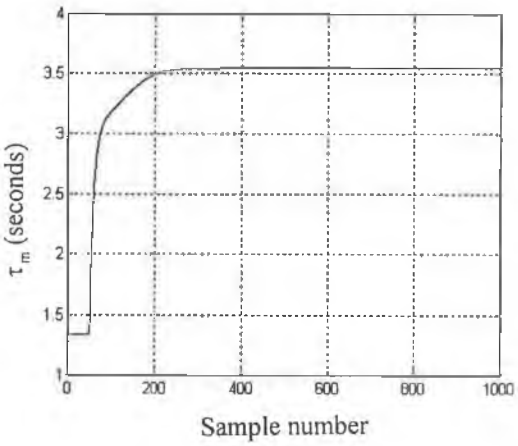


Figure 4.12: Unit step response of the process and the FOLPD model

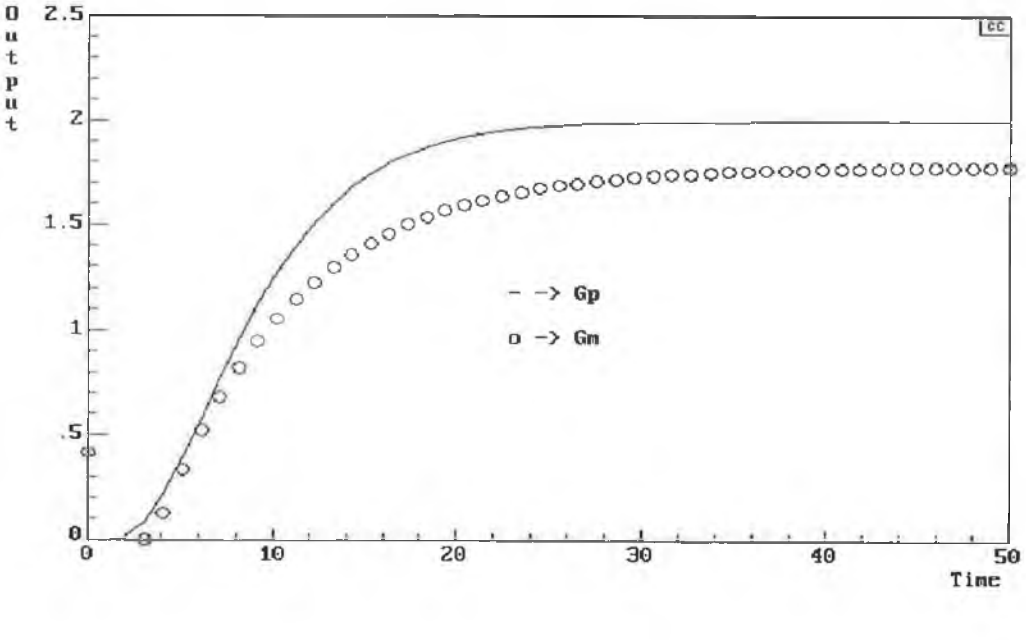
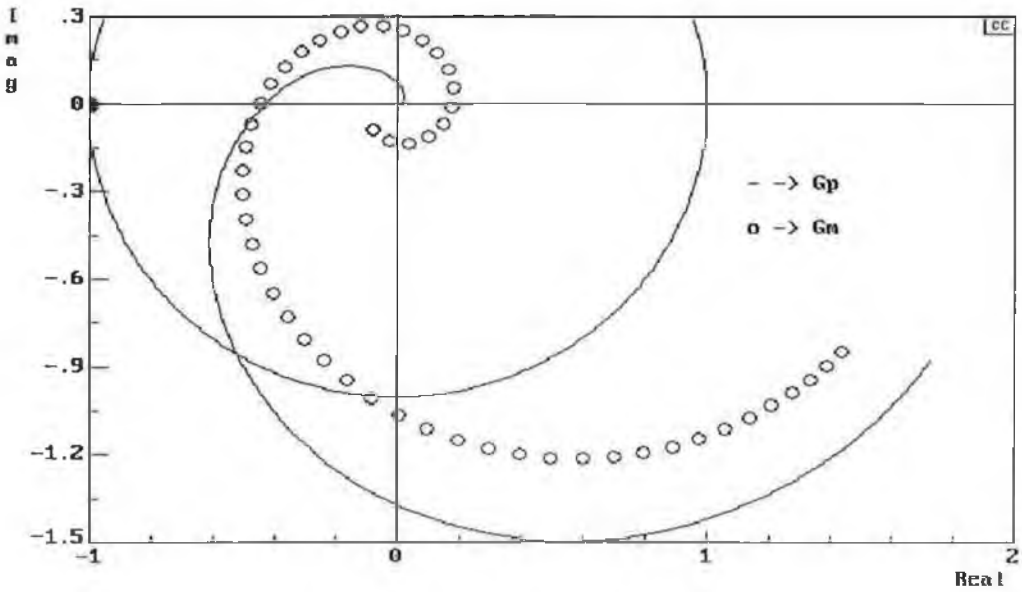


Figure 4.13: Polar plot of the process and the FOLPD model



4.3.3.2 SOSPD model parameter estimation (with no zero)

The transfer function of the model is defined as

$$G_m(s) = \frac{K_m e^{-s\tau_m}}{1 + a_{1m}s + a_{2m}s^2} \tag{4.132}$$

From equations (4.73), (4.78), (4.79), (4.84) and (4.85), the parameters of the model are calculated analytically as follows:

$$K_m = \sqrt{\frac{\left(\omega_2^2\omega_1^2 + \omega_3^4\right)\left(\omega_2^2 - \omega_1^2\right) - \omega_3^2\left(\omega_2^4 - \omega_1^4\right)}{\dfrac{\omega_2^2\omega_1^2\left(\omega_2^2 - \omega_1^2\right)}{\left|G_p(j\omega_3)\right|^2} - \dfrac{\omega_3^2\omega_2^2\left(\omega_2^2 - \omega_3^2\right)}{\left|G_p(j\omega_1)\right|^2} + \dfrac{\omega_1^2\omega_3^2(\omega_1^2 - \omega_3^2)}{\left|G_p(j\omega_2)\right|^2}}} \tag{4.133}$$

$$a_{2m} = \sqrt{\frac{K_m^2\left(\dfrac{\omega_2}{\left|G_p(j\omega_1)\right|}\right)^2 - K_m^2\left(\dfrac{\omega_1}{\left|G_p(j\omega_2)\right|}\right)^2 + \left(\omega_1^2 - \omega_2^2\right)}{\omega_1^2\omega_2^2\left(\omega_1^2 - \omega_2^2\right)}} \tag{4.134}$$

$$a_{1m} = \sqrt{\frac{\left(\frac{K_m}{|G_p(j\omega)|}\right)^2 - \left(1 - a_{2m}\omega^2\right)^2}{\omega^2}}, \quad \omega = \omega_1, \omega_2 \text{ or } \omega_3 \tag{4.135}$$

$$\tau_m = \frac{1}{\omega}\left[-\phi_p(j\omega) - \tan^{-1}\left(\frac{a_{1m}\omega}{1 - a_{2m}\omega^2}\right)\right], \quad \omega = \omega_1, \omega_2 \text{ or } \omega_3 \tag{4.136}$$

with ω_1, ω_2 and ω_3 being appropriate frequency values.

Numerical evaluation of the parameter estimates calculated from equations (4.133) to (4.136) show that the sensitivity of the model parameters calculated to changes in the process magnitude and phase values recorded is reduced, under the following conditions:

(a) The gain, K_m, is calculated from three magnitude values that span at least a decade of frequency.

(b) The parameter a_{2m} is calculated from magnitudes recorded at least a decade apart in frequency.

(c) The parameter a_{1m} is calculated at frequencies when $0.25K_m < |G_p(j\omega)| < 0.75K_m$.

(d) The time delay, τ_m, is calculated at frequencies when $|G_p(j\omega)| < 0.5K_m$.

These rules of thumb are broadly similar to those determined when calculating the parameters of a FOLPD model using the analytical method.

The analytical estimates of the SOSPD model parameters are calculated for each of the processes indicated in Cases A to G, using equations (4.133) to (4.136), based on the process frequency response values also used to calculate the FOLPD model parameters. Average values of the parameters are calculated over a number of points of the frequency response (to improve the robustness of the estimates). The frequency response data values used in equations (4.133) to (4.136) were chosen to conform with the rules of thumb defined that reduce the sensitivity of the estimates to the magnitude values recorded. The model parameters were estimated under both noise free conditions and the ±10% noise condition. The estimates of the model parameters determined are summarised in Table 4.3.

Table 4.3: SOSPD model parameter estimates, calculated using the analytical method

	No Noise				±10% Noise			
	K_m	a_{1m}	a_{2m}	τ_m	K_m	a_{1m}	a_{2m}	τ_m
Case A	2.00	1.01	0.02	0.49	2.19	1.24	0.10	0.42
Case B	1.99	4.49	4.57	0.99	2.24	5.36	4.98	1.04
Case C	1.97	7.20	19.72	1.70	2.19	8.85	22.45	1.72
Case D	2.00	11.15	46.83	7.80	2.20	11.88	57.46	7.47
Case E	2.00	0.99	1.00	1.00	2.19	1.03	1.19	1.02
Case F	2.00	6.30	8.15	1.00	2.22	7.34	9.42	1.01
Case G	1.99	6.28	8.33	4.37	2.28	7.89	7.67	4.79

The use of the gradient algorithm (equation (4.103)) requires that the cost function be unimodal with respect to a_{1m} and a_{2m}. Equations (4.113) and (4.114), which define the unimodality condition, reduce (for this application) to

$$\frac{\partial^2 J}{\partial a_{1m}^2} = 0.5\sum_{n=1}^{3}\{\frac{\omega_n^4 K_m^2 a_{1m}^2}{|G_p(j\omega_n)|\left[\left(1-a_{2m}\omega_n^2\right)^2+(a_{1m}\omega_n)^2\right]^3}+\frac{\omega_n(1-a_{2m}\omega_n^2)^2}{\left[\left(1-a_{2m}\omega_n^2\right)^2+(a_{1m}\omega_n)^2\right]^2}\}+$$

$$0.5\sum_{n=1}^{3}\{\frac{K_m\omega_n^2[2\omega_n^2 a_{1m}^2-(1-a_{2m}\omega_n^2)^2]}{|G_p(j\omega_n)|\left[\left(1-a_{2m}\omega_n^2\right)^2+\left(a_{1m}\omega_n\right)^2\right]^{2.5}}\left[\frac{K_m}{\sqrt{\left(1-a_{2m}\omega_n^2\right)^2+(a_{1m}\omega_n)^2}}-|G_p(j\omega_n)|\right]\}$$

$$-0.5\sum_{n=1}^{3}\{\frac{2\omega_n^2 a_{1m}\left(1-a_{2m}\omega_n^2\right)}{\left[\left(1-a_{2m}\omega_n^2\right)^2+(a_{1m}\omega_n)^2\right]^2}\left[\tan^{-1}\frac{\omega_n a_{1m}}{1-a_{2m}\omega_n^2}+\omega_n\tau_m+\phi_p(j\omega_n)\right]\} > 0 \qquad (4.137)$$

and

$$\frac{\partial^2 J}{\partial a_{2m}^2} = 0.5\sum_{n=1}^{3}\{\frac{\omega_n^4 K_m^2(1-a_{2m}\omega_n^2)^2}{|G_p(j\omega_n)|\left[\left(1-a_{2m}\omega_n^2\right)^2+(a_{1m}\omega_n)^2\right]^3}+\frac{a_{1m}^2\omega_n^5}{\left[\left(1-a_{2m}\omega_n^2\right)^2+(a_{1m}\omega_n)^2\right]^2}\}+$$

$$0.5\sum_{n=1}^{3}\frac{K_m\omega_n^4[2(1-a_{2m}\omega_n^2)^2-\left(a_{1m}\omega_n\right)^2]}{|G_p(j\omega_n)|\left[\left(1-a_{2m}\omega_n^2\right)^2+(a_{1m}\omega_n)^2\right]^{2.5}}\left[\frac{K_m}{\sqrt{\left(1-a_{2m}\omega_n^2\right)^2+(a_{1m}\omega_n)^2}}-|G_p(j\omega_n)|\right]$$

$$+0.5\sum_{n=1}^{3}\frac{2\omega_n^4 a_{1m}\left(1-a_{2m}\omega_n^2\right)}{\left[\left(1-a_{2m}\omega_n^2\right)^2+(a_{1m}\omega_n)^2\right]^2}\left[\tan^{-1}\frac{\omega_n a_{1m}}{1-a_{2m}\omega_n^2}+\omega_n\tau_m+\phi_p(j\omega_n)\right]>0 \qquad (4.138)$$

Numerical evaluation of equations (4.137) and (4.138) revealed that the initial estimates of the parameter values to be used in the gradient approach should be as defined in either Option 3 or Option 4, to facilitate unimodality of the cost function with respect to the variation in a_{1m} and a_{2m}.

Option 3:

$$0.83K_m(\text{optimum}) \le K_m(\text{analytical}) \le 1.17K_m(\text{optimum}) \qquad (4.139)$$

$$0.75a_{1m}(\text{optimum}) \le a_{1m}(\text{analytical}) \le 1.5a_{1m}(\text{optimum}) \qquad (4.140)$$

$$0.5a_{2m}(\text{optimum}) \le a_{2m}(\text{analytical}) \le 1.75a_{2m}(\text{optimum}) \qquad (4.141)$$

$$0.83\tau_m(\text{optimum}) \le \tau_m(\text{analytical}) \le 1.17\tau_m(\text{optimum}) \qquad (4.142)$$

with $a_{1m}(\text{optimum})$ and $a_{2m}(\text{optimum})$ being the least squares value of a_{1m} and a_{2m} calculated using the gradient method.

Option 4:

$$K_m(\text{gradient}-\text{initial}) = 1.5K_m(\text{analytical}) \qquad (4.143)$$

$$0.75a_{1m}(\text{optimum}) \le a_{1m}(\text{analytical}) \le 1.75a_{1m}(\text{optimum}) \qquad (4.144)$$

$$0.5a_{2m}(\text{optimum}) \le a_{2m}(\text{analytical}) \le 1.75a_{2m}(\text{optimum}) \qquad (4.145)$$

$$\tau_m(\text{gradient}-\text{initial}) = 0.5\tau_m(\text{analytical}) \qquad (4.146)$$

For the estimation of the parameters of a SOSPD model, the specifications in Options 3 and 4 are broadly similar; as with Options 1 and 2, both Options 3 and 4 describe worst case conditions. The analytical estimates of the parameters are first calculated (using equations (4.133) to (4.136)); these may then be used as the initial values for the gradient estimates of the parameters (assuming Option 3 is taken).

Convergence to the optimum values of the parameter estimates using the gradient method is of course not guaranteed (as in the case when the parameters of a FOLPD model are being estimated). Simulation results show convergence to fixed SOSPD parameter estimates for all of the processes under discussion, under both the noise free and ±10% noise conditions considered; the convergence is, generally speaking, slower than when the parameters of a FOLPD model are being estimated. These simulation results are summarised in Table 4.4.

Table 4.4: SOSPD model parameter estimates, calculated using the gradient method

	No Noise				±10% Noise			
	K_m	a_{1m}	a_{2m}	τ_m	K_m	a_{1m}	a_{2m}	τ_m
Case A	2.00	1.02	0.02	0.48	2.02	1.05	0.04	0.47
Case B	2.01	4.54	4.51	1.00	1.99	4.48	4.30	1.05
Case C	1.94	7.34	16.76	1.89	1.91	7.22	15.95	2.00
Case D	1.87	10.35	38.20	8.25	1.78	9.66	30.28	9.08
Case E	2.00	1.00	1.00	1.00	2.00	1.01	0.98	1.04
Case F	2.01	6.32	8.25	0.99	2.03	6.42	7.90	1.04
Case G	2.22	7.67	16.01	3.73	2.22	7.54	16.14	3.79

The open loop step and frequency response of the SOSPD models, formed from the parameter estimates in Table 4.4, under both the noise free and ±10% noise conditions, were compared with the responses of the corresponding processes. These results, determined using Program CC, show that when the process is itself of a FOLPD or SOSPD structure, excellent fitting was achieved (as may be deduced from the results provided for Cases A, B and E in Table 4.4) though when the process was not of this structure, an expected deviation between the process and the model in both the time domain and the frequency domain was observed. This deviation was, however, far less than that seen when a FOLPD model was estimated (particularly in the frequency domain), which indicates that if accurate fitting of the model to the process in the frequency domain is required, then the estimation of a SOSPD model of the process is more appropriate than a FOLPD model. However, a caveat is that better initial estimates of the gradient model parameters are required, as the worst case conditions for convergence are tighter when estimating the parameters of a SOSPD model compared to estimating the parameters of a FOLPD model, as a comparison of Options 1 and 4 or Options 2 and 3 will reveal. A further disadvantage of estimating the parameters of a SOSPD model is the slower convergence rate of the parameter estimates (though the speed of convergence of the parameter estimates may be altered by varying the learning rate, μ (equation (4.103)). As is the case when estimating a FOLPD model, the acceptability of the fitting depends on the use to which the process model is applied.

One representative simulation result is provided in Figures 4.14 to 4.29, for the process in Case C, under the ±10% noise condition. The model parameter values calculated analytically in this case are $K_m = 2.19$, $a_{1m} = 8.85$, $a_{2m} = 22.45$ and $\tau_m = 1.72$ seconds (Table 4.3).

Figures 4.14 to 4.23 show that, for the simulation taken, a wide range of initial parameter values is possible (O = points where the appropriate second partial derivative is less than zero, • = points where the appropriate second partial derivative is greater than zero and [] = approximate allowed range of the analytical parameter estimates of a_{1m} and a_{2m}, over all values of K_m and τ_m taken (Option 3); Option 3 describes a conservative bound for this particular simulation). Figures 4.14 to 4.18 record when $\partial^2 J/\partial {a_{1m}}^2 < 0$; Figures 4.19 to 4.23 record when $\partial^2 J/\partial {a_{2m}}^2 < 0$.

Figures 4.24 to 4.27 show the convergence of the initial parameter values to final values within 500 samples, using the gradient method. The final values of $K_m = 1.91$, $a_{1m} = 7.22$, $a_{2m} = 15.95$ and $\tau_m = 2.00$ seconds (Table 4.4) mean that the guidelines suggested in Option 3 are fulfilled.

Figures 4.28 and 4.29 show the step response and frequency response of the process and model together (using Program CC). The fitting of the process to the model in both domains is excellent, and is better than if a FOLPD model is estimated (Figures 4.12 and 4.13).

Fig. 4.14: $K_m = 1.58, \tau_m = 1.67$

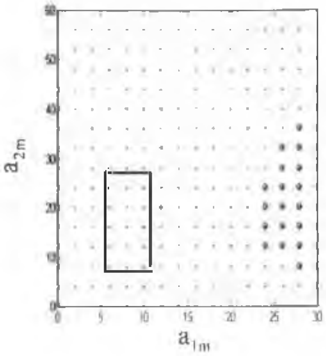


Fig. 4.15: $K_m = 1.58, \tau_m = 2.33$

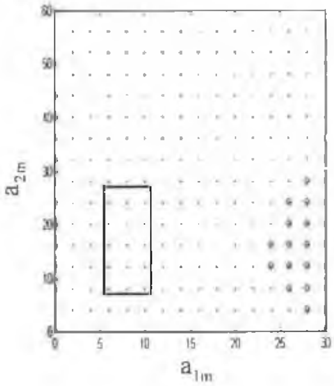


Fig. 4.16: $K_m = 2.23, \tau_m = 1.67$

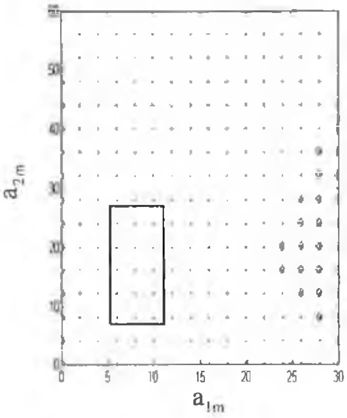


Fig. 4.17: $K_m = 2.23, \tau_m = 2.33$

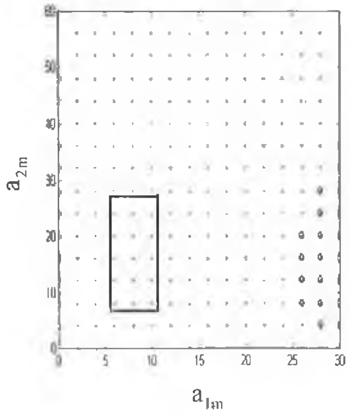


Fig. 4.18: $K_m = 1.91, \tau_m = 2.00$

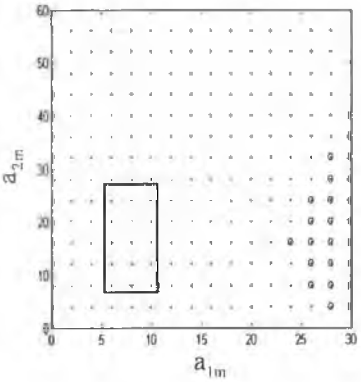


Fig. 4.19: $K_m = 1.91, \tau_m = 2.00$

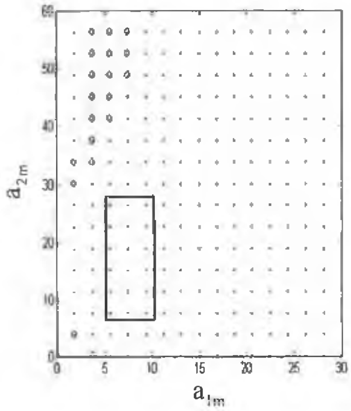


Fig. 4.20: $K_m = 1.58, \tau_m = 1.67$

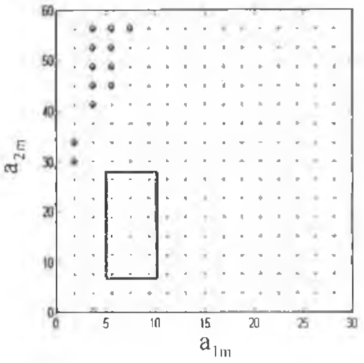


Fig. 4.21: $K_m = 1.58, \tau_m = 2.33$

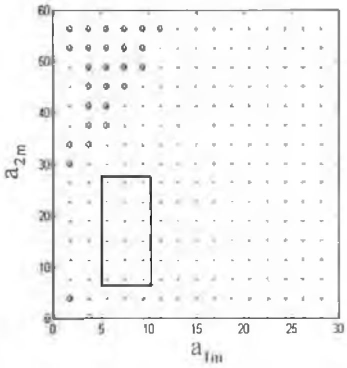


Fig. 4.22: $K_m = 2.23, \tau_m = 1.67$

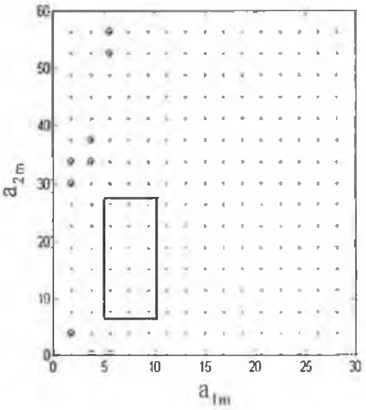


Fig. 4.23: $K_m = 2.23, \tau_m = 2.33$

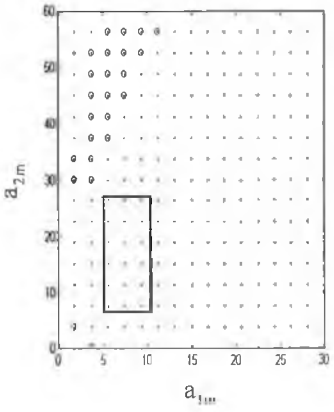


Figure 4.24: $K_m - \mu = 1.0$

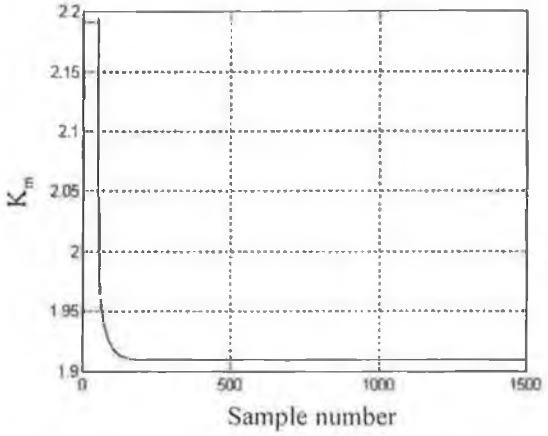


Figure 4.25: $a_{1m} - \mu = 10.0$

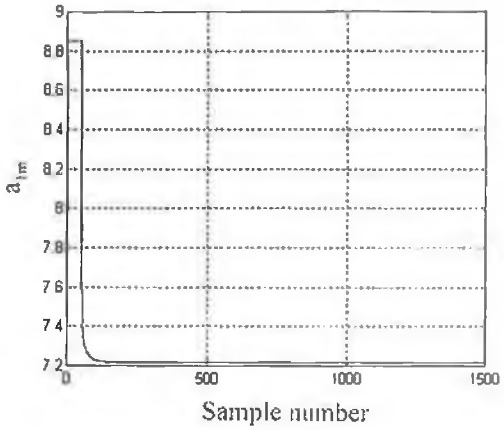


Figure 4.26: $a_{2m} - \mu = 10.0$

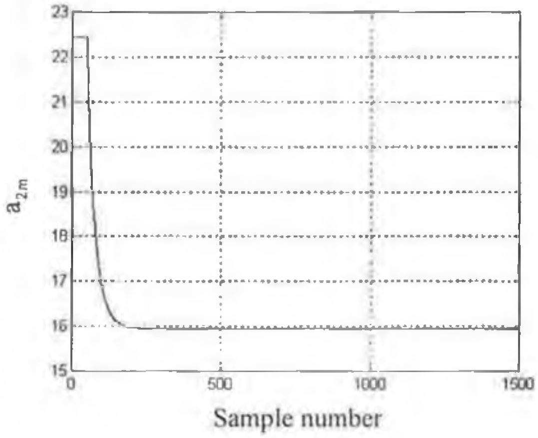


Figure 4.27: $\tau_m - \mu = 0.1$

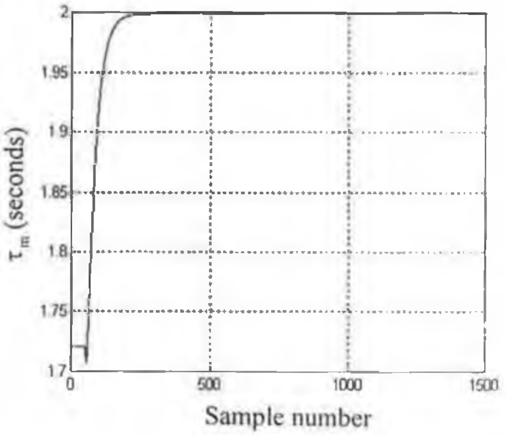


Figure 4.28: Unit step response of the process and the SOSPD model

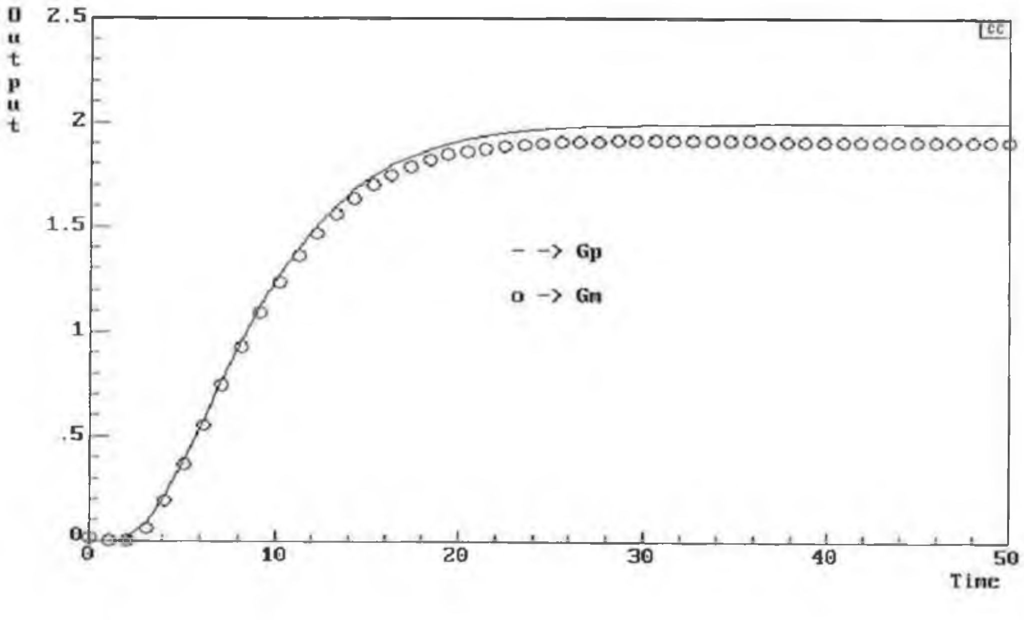
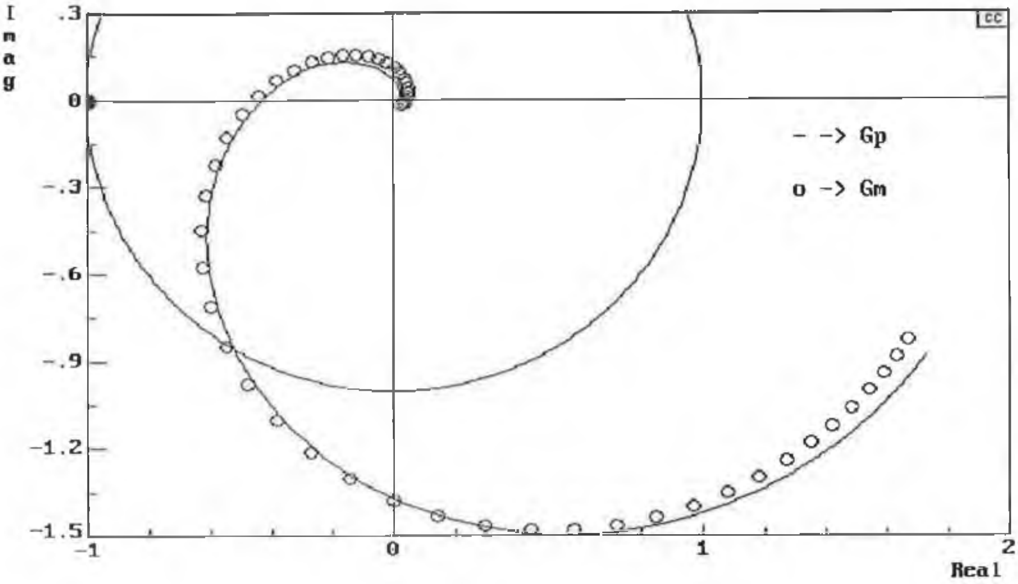


Figure 4.29: Polar plot of the process and the SOSPD model



4.3.3.3 Estimating the parameters of a third order model (with time delay) and no zeroes

Separate initial estimates of the parameters are calculated using the analytical formulae that estimate the parameters of a general third order model (with no zeroes), and using the analytical formulae that estimate the parameters of a third order model with a multiple pole (and with no zeroes). Using the former approach, the transfer function of the model is

$$G_m(s) = \frac{K_m e^{-s\tau_m}}{1 + a_{1m}s + a_{2m}s^2 + a_{3m}s^3} \qquad (4.147)$$

From equation (4.83), the non-delay parameters are calculated by solving the following relationship:

$$\begin{bmatrix} |G_p(j\omega_1)|^2 \\ |G_p(j\omega_2)|^2 \\ |G_p j\omega_3)|^2 \\ |G_p(j\omega_4)|^2 \end{bmatrix} = \begin{bmatrix} 1 & -\omega_1^2|G_p(j\omega_1)|^2 & -\omega_1^4|G_p(j\omega_1)|^2 & -\omega_1^6|G_p(j\omega_1)|^2 \\ 1 & -\omega_2^2|G_p(j\omega_2)|^2 & -\omega_2^4|G_p(j\omega_2)|^2 & -\omega_2^6|G_p(j\omega_2)|^2 \\ 1 & -\omega_3^2|G_p(j\omega_3)|^2 & -\omega_3^4|G_p(j\omega_3)|^2 & -\omega_3^6|G_p(j\omega_3)|^2 \\ 1 & -\omega_4^2|G_p(j\omega_4)|^2 & -\omega_4^4|G_p(j\omega_4)|^2 & -\omega_4^6|G_p(j\omega_4)|^2 \end{bmatrix} \begin{bmatrix} K_m^2 \\ c_{1m} \\ c_{2m} \\ c_{3m} \end{bmatrix}$$

(4.148)

Applying equations (4.78) to (4.81), it may be deduced that

$$a_{3m} = \sqrt{c_{3m}} \tag{4.149}$$

with a_{2m} solved numerically from the following equation:

$$a_{2m}^4 - 2c_{2m}a_{2m}^2 - 8c_{3m}a_{2m} + (c_{2m}^2 - 4c_{1m}c_{3m}) = 0 \tag{4.150}$$

and with

$$a_{1m} = \sqrt{c_{1m} + 2a_{2m}} \tag{4.151}$$

The time delay is calculated from equation (4.85), as follows:

$$\tau_m = \frac{1}{\omega}[-\phi_p(j\omega) - \tan^{-1}((a_{1m}\omega - a_{3m}\omega^3)/(1 - a_{2m}\omega^2))] \tag{4.152}$$

with $\omega = \omega_1, \omega_2, \omega_3$ or ω_4. The alternative to this procedure is to estimate the model parameters, assuming a repeated pole on the denominator of the transfer function i.e.

$$G_{m1}(s) = \frac{K_{m1}e^{-s\tau_{m1}}}{(1 + sT_{m1})^3} \tag{4.153}$$

From equations (4.89), (4.90) and (4.91), the parameters K_{m1}, T_{m1} and τ_{m1}, respectively, are calculated as follows:

$$K_{m1} = \frac{|G_p(j\omega_1)||G_p(j\omega_2)|(\omega_2{}^2 - \omega_1{}^2)^{1.5}}{\left[\left(\omega_2|G_p(j\omega_2)|^{0.333}\right)^2 - \left(\omega_1|G_p(j\omega_1)|^{0.333}\right)^2\right]^{1.5}} \tag{4.154}$$

$$T_{m1} = \frac{1}{\omega}\sqrt{\left[\frac{K_{m1}}{|G_p(j\omega)|}\right]^{0.667} - 1}, \quad \omega = \omega_1 \text{ or } \omega_2 \tag{4.155}$$

and

$$\tau_{m1} = \frac{1}{\omega}\left[-\phi_p(j\omega) - 3\tan^{-1}(\omega T_{m1})\right], \quad \omega = \omega_1 \text{ or } \omega_2 \qquad (4.156)$$

The denominator parameters a_{1m}, a_{2m}, a_{3m} are then calculated in a straightforward manner.

The model parameters are calculated under the noise free and ± 10% noise conditions. The analytical estimates of the parameters (in equations (4.148) to (4.152)) are calculated following broadly the conditions defined for the analytical calculation of the parameters in the SOSPD case i.e. the parameters K_m, a_{3m}, a_{2m} and a_{1m} are calculated from three magnitude values that span at least a decade of frequency and the time delay, τ_m, is calculated at frequency values when $|G_p(j\omega)| < 0.5K_m$. Detailed simulation results to determine the robustness of the parameter estimates to changes in the magnitude values recorded would need to be carried out if refinements in these rules of thumb were judged desirable. The parameters in equations (4.154) to (4.156) are calculated based on the conditions defined for reduced sensitivity of the parameters of a FOLPD model to process magnitude values recorded.

The model parameter estimates calculated using equations (4.148) to (4.152), for Cases A to G, are summarised in Table 4.5, with the model parameter estimates calculated using equations (4.154) to (4.156) summarised in Table 4.6.

Table 4.5: Model parameters calculated using the analytical method

	No Noise					±10% Noise				
	K_m	a_{1m}	a_{2m}	a_{3m}	τ_m	K_m	a_{1m}	a_{2m}	a_{3m}	τ_m
Case A	2.00	1.09	0.10	0.00	0.41	2.33	2.20	0.85	0.16	0.05
Case B	1.99	4.44	4.33	0.23	0.99	2.22	5.00	5.54	0.00	0.95
Case C	1.99	8.38	22.1	16.3	1.13	2.36	13.2	38.5	56.2	0.60
Case D	2.00	13.5	66.1	123	5.73	2.37	29.3	200	679	2.71
Case E	2.00	1.31	0.85	0.09	1.04	2.35	2.55	1.98	0.77	0.65
Case F	2.00	6.29	8.54	0.00	0.96	2.33	9.92	18.3	16.8	0.29
Case G	1.99	6.34	9.21	0.00	4.25	2.16	7.55	22.9	0.00	3.32

Table 4.6: Model parameters calculated using the analytical method (repeated pole)

	No Noise					±10% Noise				
	K_m	a_{1m}	a_{2m}	a_{3m}	τ_m	K_m	a_{1m}	a_{2m}	a_{3m}	τ_m
Case A	1.17	0.65	0.14	0.01	0.29	1.29	0.73	0.18	0.01	0.27
Case B	1.94	4.85	7.83	4.22	0.25	2.15	6.43	13.8	9.84	0.00
Case C	1.98	8.57	24.5	23.3	0.79	2.18	10.6	37.4	44.1	0.24
Case D	2.00	13.1	57.5	84.0	6.58	2.21	14.6	71.4	116	6.14
Case E	2.08	1.54	0.79	0.13	1.12	2.30	1.57	0.83	0.14	1.18
Case F	1.75	6.21	12.9	8.88	0.12	1.92	6.63	14.7	10.8	0.08
Case G	1.97	7.01	16.4	12.8	3.02	2.17	7.74	20.0	17.2	2.98

Comparing the results for Cases A, B, C and E in Table 4.5 and 4.6, for the noise free case, it is evident that the values calculated using the more general model structure are closer to the actual process values, then are the values calculated using the model structure with a repeated pole. The trade-off is the much higher computational cost associated with the calculation of the model parameters in the former structure, particularly if any of the parameters must be calculated numerically (equation (4.150)).

As the model order increases, the evaluation of the allowed range of estimates of the analytical parameter values, for successful implementation of the gradient method, becomes more time-consuming. Since a requirement for unimodality is that the second partial derivative of the cost function with respect to the denominator parameter values be greater than zero, it was decided to increase the initial estimate of the model gain, and decrease the initial estimate of the model time delay, if any of the second partial derivatives of the cost function with respect to the denominator parameter values were less than zero (it was felt that increasing the model gain, and decreasing the model time delay would be more likely to facilitate unimodality of the cost function, following the example of the FOLPD model parameter estimation strategy). The model gain is increased, and the model time delay is reduced, until a set of model parameter values is reached when all of the relevant second partial derivatives are greater than zero.

Simulation results show convergence of the model parameters to their optimum values, using the gradient method, in some but not all simulated processes, when either a general third order model or a third order model with a repeated pole was used. These simulation results are summarised in Tables 4.7 and 4.8, respectively.

Table 4.7: Parameters calculated using the gradient method - general model

	No Noise					±10% Noise				
	K_m	a_{1m}	a_{2m}	a_{3m}	τ_m	K_m	a_{1m}	a_{2m}	a_{3m}	τ_m
Case A	2.00	1.09	0.09	0.00	0.41	2.29	1.38	0.31	0.02	0.18
Case B	2.01	4.54	4.54	0.03	0.99	2.00	4.65	4.95	0.70	0.90
Case C	2.02	8.00	33.7	11.0	0.09	2.00	7.94	34.3	10.6	0.01
Case D	1.97	13.7	66.3	147	5.41	2.19	15.0	139	227	0.76
Case E	2.00	1.00	1.00	0.00	1.00	2.14	1.36	1.39	0.30	0.57
Case F	2.01	6.33	8.36	0.14	0.97	2.05	6.70	16.1	2.50	0.01
Case G	2.22	7.67	15.5	0.00	3.77	2.21	7.44	12.2	0.00	4.17

Table 4.8: Parameters calculated using the gradient method - repeated pole model

	No Noise					±10% Noise				
	K_m	a_{1m}	a_{2m}	a_{3m}	τ_m	K_m	a_{1m}	a_{2m}	a_{3m}	τ_m
Case A	2.00	1.10	0.10	0.00	0.41	2.06	1.25	0.19	0.02	0.31
Case B	2.08	4.95	9.80	1.80	0.01	2.06	4.90	10.3	1.80	0.01
Case C	2.13	8.62	35.0	11.9	0.00	2.09	8.40	34.7	12.0	0.03
Case D	1.97	13.6	65.7	144	5.46	1.89	13.3	56.5	132	6.18
Case E	2.00	1.00	1.00	0.00	1.00	2.00	1.01	0.98	0.00	1.04
Case F	2.04	6.57	15.7	2.5	0.00	2.08	6.70	16.2	2.60	0.02
Case G	2.22	7.66	15.5	0.00	3.76	2.21	7.46	12.4	0.30	4.14

Both sets of simulation results show that in some cases, convergence of the parameters, using the gradient method, to their optimum values (or values close to the optimum values) is facilitated. This may be seen directly in Table 4.7, in the following cases:

(1) Case A, no noise

(2) Case B, no noise and ±10% noise and

(3) Case E, no noise.

In addition, supplementary simulation results (O'Dwyer (1996h)) demonstrate good fitting between the process and the models, in Program CC, for the following cases in Table 4.7:

(1) Case D, no noise

(2) Case F, no noise

(3) Case G, no noise and ±10% noise.

The convergence of the parameters may be seen directly in Table 4.8, in the following cases:

(1) Case A, no noise and ±10% noise

(2) Case E, no noise and ±10% noise

In addition, supplementary simulation results (O'Dwyer (1996h)) demonstrate good fitting between the process and the models, in Program CC, for the following cases in Table 4.8:

(1) Case D, no noise and ±10% noise

(2) Case G, no noise and ±10% noise.

However, it is disappointing that the optimum parameters of a third order model plus time delay (Case C) were not estimated using the gradient method, for the algorithms defined. Thus, the tactic of increasing the model gain, and decreasing the model time delay, to ensure that all of the second partial derivatives are greater than zero does not facilitate optimum parameter estimation in all cases. A more sophisticated strategy may need to be employed to check that, as all of the parameters vary, the second partial derivatives remain positive at each iteration. This suggestion would, however, involve a large number of calculations at each sample time. Alternatively, initial values of other parameters could be modified to try to ensure positive second partial derivatives; unfortunately, there would be no guarantee that the second partial derivatives would remain positive as the parameters changed. The conclusion from this discussion is that it is not a straightforward matter to facilitate convergence of the model parameters to their optimum values if any of the second partial derivatives are less than zero; under these circumstances, it may be better to estimate the parameters of a less complex model structure.

4.3.3.4 Estimating the parameters of a second order model (with time delay) and one zero

The transfer function of the model is defined as

$$G_m(s) = \frac{K_m(1 + b_{1m}s)e^{-s\tau_m}}{1 + a_{1m}s + a_{2m}s^2} \qquad (4.157)$$

From equations (4.73), (4.74), (4.78), (4.79), (4.84) and (4.85), the parameters K_m, a_{1m}, a_{2m}, b_{1m} and τ_m are calculated as follows:

$$K_m = \sqrt{d_{0m}} \qquad (4.158)$$

$$b_{1m} = \sqrt{d_{1m}}/K_m \qquad (4.159)$$

$$a_{2m} = \sqrt{c_{2m}} \qquad (4.160)$$

$$a_{1m} = \sqrt{c_{1m} + 2a_{2m}} \qquad (4.161)$$

and

$$\tau_m = \frac{1}{\omega}[-\phi_p(j\omega) + \tan^{-1}(b_{1m}\omega) - \tan^{-1}(a_{1m}\omega/(1 - a_{2m}\omega^2))] \qquad (4.162)$$

with $\omega = \omega_1, \omega_2, \omega_3$ or ω_4. The model parameters are calculated under the noise free and $\pm 10\%$ noise conditions. The analytical estimates of the parameters (in equations (4.158) to (4.162)) are calculated following broadly the conditions defined for the analytical calculation of the parameters in the SOSPD case i.e. the parameters K_m, b_{1m}, a_{2m} and a_{1m} are calculated from three magnitude values that span at least a decade of frequency and the time delay, τ_m, is calculated at frequency values when $|G_p(j\omega)| < 0.5\tau_m$. Detailed simulation results to determine the robustness of the parameter estimates to changes in the magnitude values recorded would need to be carried out if refinements in these rules of thumb were judged desirable. The model parameters calculated using equations (4.158) to (4.162) are summarised in Table 4.9.

Table 4.9: Model parameters calculated using the analytical method

	No Noise					±10% Noise				
	K_m	a_{1m}	a_{2m}	b_{1m}	τ_m	K_m	a_{1m}	a_{2m}	b_{1m}	τ_m
Case A	2.00	0.64	0.00	0.00	0.54	2.33	2.28	0.70	0.73	0.56
Case B	1.99	4.49	4.57	0.00	0.99	2.21	5.11	6.36	0.52	1.35
Case C	1.99	7.36	14.5	0.00	2.07	2.18	7.65	18.9	0.00	1.84
Case D	2.00	10.4	34.6	0.00	8.59	2.28	9.07	0.00	0.00	12.2
Case E	2.00	1.40	0.99	0.00	1.09	2.34	1.36	0.00	0.00	1.78
Case F	2.00	6.32	8.71	0.35	1.28	2.32	7.39	0.00	0.00	2.00
Case G	1.99	6.50	9.96	0.78	4.91	2.18	3.43	0.00	0.00	6.48

The analytical parameters calculated were used as the initial values for the gradient estimates of the parameters. In all cases, convergence to final values of the parameters was achieved. The simulation results obtained are summarised in Table 4.10.

Table 4.10: Model parameters calculated using the gradient method

	No Noise					±10% Noise				
	K_m	a_{1m}	a_{2m}	b_{1m}	τ_m	K_m	a_{1m}	a_{2m}	b_{1m}	τ_m
Case A	2.00	1.00	0.00	0.01	0.51	2.04	2.19	1.12	1.12	0.51
Case B	2.01	4.54	4.51	0.00	1.00	1.98	4.49	4.62	0.28	1.28
Case C	1.94	7.34	16.8	0.00	1.90	1.91	7.22	15.9	0.00	2.00
Case D	1.87	10.4	38.2	0.00	8.25	1.78	9.66	30.3	0.00	9.08
Case E	2.00	1.00	1.00	0.00	1.03	2.00	1.01	0.98	0.00	1.03
Case F	2.00	6.32	8.39	0.19	1.17	2.03	6.42	7.62	0.00	1.06
Case G	2.20	7.74	13.7	0.85	4.75	2.16	6.46	0.00	0.05	5.78

A comparison of the results in Table 4.4 and Table 4.10 shows that, in many cases, very similar results are obtained when the processes are modelled by a SOSPD model without a zero (Table 4.4) and with a zero (Table 4.10). This is not surprising, as many of the processes do not contain a zero. In Case G, under the ±10% noise condition, the results in Table 4.10 may be more directly compared with the results in Table 4.2 (when the processes are modelled by a FOLPD model). It is possible to identify an overdetermined model if a zero is estimated; the results for Case A under the ±10% noise condition in Table 4.10 shows that the model identified has a common factor (approximately) on the numerator and denominator terms.

Overall, there does not appear to be a significant benefit in estimating the parameters of a model with zeroes as opposed to estimating the parameters of a model without zeroes. In specific terms, the following reasons suggest why it may be more appropriate to estimate a process model without zeroes:

(a) There is an increased computational burden involved in estimating the parameters of a model with zeroes.

(b) It is likely that there will be an increased difficulty in calculating sufficiently good parameter estimates using the analytical approach (if the parameters of a model with zeroes are being estimated), so that a reasonable possibility of convergence to the optimum parameter estimates using the gradient approach exists.

(c) There is reasonably good fitting, for both the noise-free and noisy cases, in both the time and frequency domains, between the processes taken, and an appropriate model without zeroes.

4.3.5 Model structure selection

The estimation of the most appropriate model of the process to use is a difficult issue. One approach is to assume that the process is adequately modelled by either a FOLPD model or a SOSPD model (with no zero). Such an assumption is frequently made in process model identification, as described in Chapter 1. The advantages and disadvantages of estimating a FOLPD model versus a SOSPD model have been debated in Section 4.3.3.2.

One simple test that may allow an indication of the preferred model to use, is to calculate the slope of the process magnitude versus frequency curve at high frequencies. This slope should be -20 dB/decade if the process is of FOLPD structure and should be -40 dB/decade if the process is of SOSPD structure. Experimentally obtained frequency response data are seldom accurate enough to exhibit a slope more negative than -40 dB/decade (Seborg *et al.* (1989)). The slope of the process magnitude versus frequency curve is given by:

$$\frac{20\log_{10}\left|G_p(j\omega_1)\right| - 20\log_{10}\left|G_p(j\omega_2)\right|}{\log_{10}(\omega_2/\omega_1)} \qquad (4.163)$$

However, the final decision on the most appropriate model structure to choose depends more on parameter convergence and on computational issues, and goodness of fit requirements.

Alternatively, the parameters of an arbitrary order model could be estimated. Some measure of the most appropriate model order to use in the estimation is necessary; one way to do this would be to calculate the cost function formed from the optimum parameters estimated (using the gradient method) as the model order is increased. Then, the value of the model order corresponding to where the cost function levels out would be the most appropriate model order to use. This procedure is computationally intensive. A variation of the strategy that is less computationally intensive would be to calculate the cost function based on the initial model parameter estimates (calculated using an analytical approach). A repeated pole model would simplify the calculations further.

It has been shown in Section 4.3.3.3 that convergence of the parameters to their optimum values using gradient methods is not always facilitated for third order process models (with time delay), due to non-unimodality of the cost function. In addition, simple means to ensure that all of the second partial derivatives are greater than zero (i.e. that the cost function is unimodal) do not always allow appropriate model parameter estimation. It is likely that the same experience would be repeated if the parameters of higher order models with time delay are to be estimated. Therefore, a strategy for the estimation of the parameters of an appropriate arbitrary order model plus time delay is summarised in Figure 4.30.

Table 4.11 summarises the results obtained when the algorithm in Figure 4.30 is applied to estimate the model order of the seven simulated processes in Cases A to G.

This test is not entirely reliable (as may be seen from Table 4.11); however, it appears to be a reasonable guide to the choice of an appropriate model order. The test has been developed assuming that the model has no zeroes; a similar test could be developed for a model with zeroes.

Figure 4.30: Flowchart summarising the algorithm for model parameter estimation

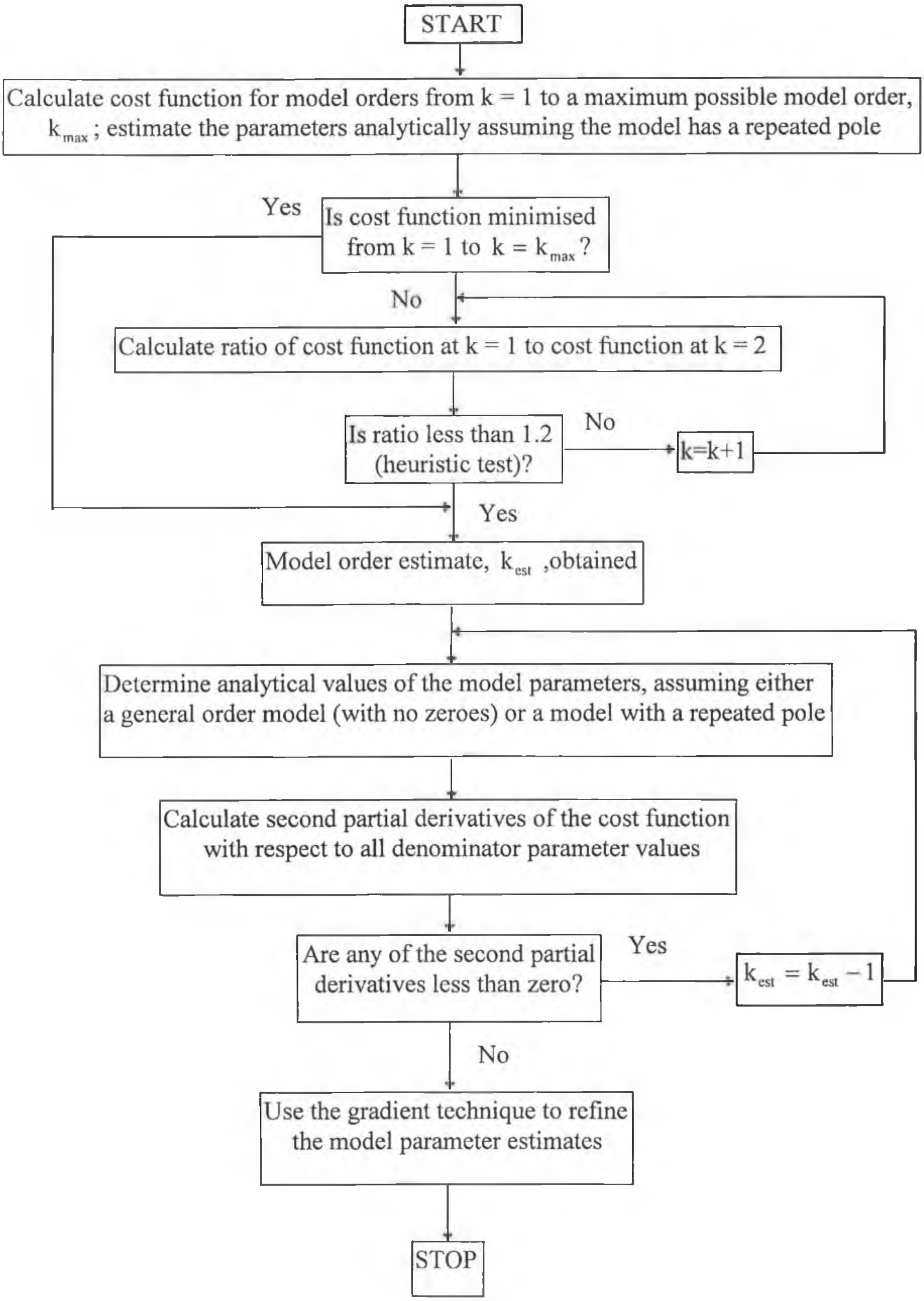


Table 4.11: Model orders estimated

Simulation	Correct Model Order	Estimated Model Order - No noise	Estimated Model Order - ± 10% Noise
Case A	1	1	1
Case B	2	2	2
Case C	3	2	2
Case D	8	6	6
Case E	2	2	2
Case F	3	2	2
Case G	3	1	2

It has already been shown that good fitting between the process and the model exists for Case C for the estimated model order value in Table 4.11, under the ±10% noise condition (Figures 4.28, 4.29). Simulation results are now presented to show the estimation of the model parameters, and the fitting between the process and the model, for Cases A, B, E, F and G. In all cases, the results are determined under the ±10% noise condition.

For case A, $G_p = 2e^{-0.5s}/(1+s)$. From Table 4.11, the parameters of a FOLPD model are estimated (Figures 4.31, 4.32 and 4.33) following the procedure in Option 1 (Section 4.3.3.1). Fitting in the time domain and in the frequency domain, using Program CC, is shown in Figures 4.34 and 4.35, respectively.

Figure 4.31: $K_m - \mu = 0.5$

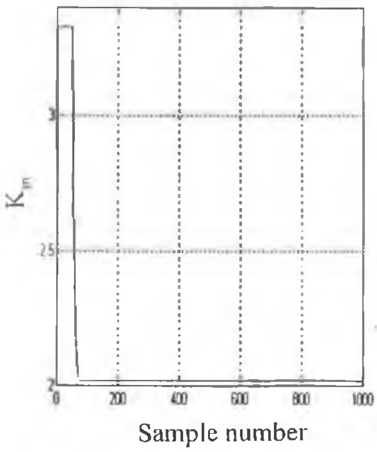


Figure 4.32: $T_m - \mu = 0.5$

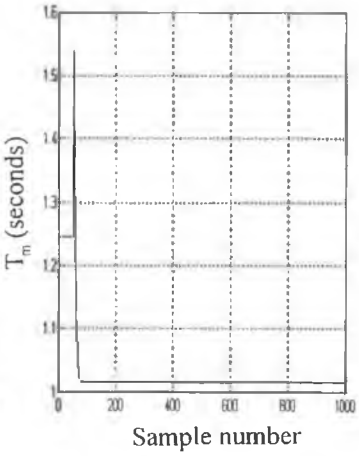


Figure 4.33: $\tau_m - \mu = 0.5$

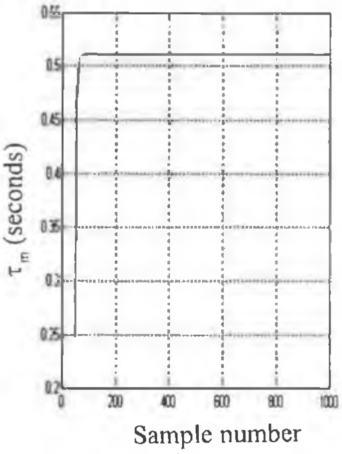


Figure 4.34: Unit step response of the process and the FOLPD model

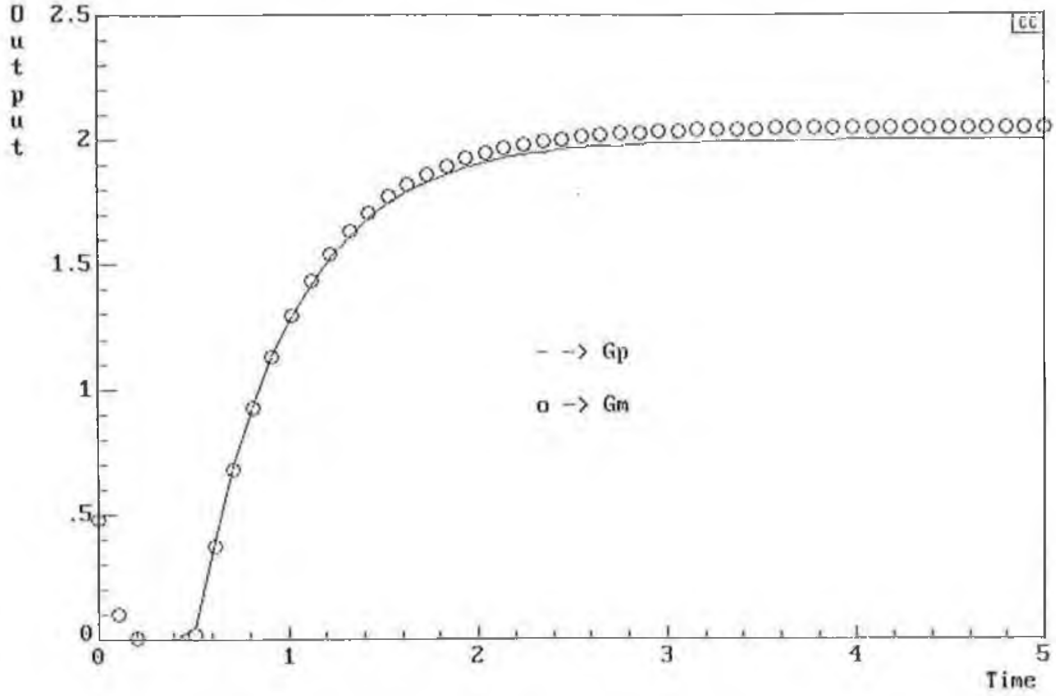
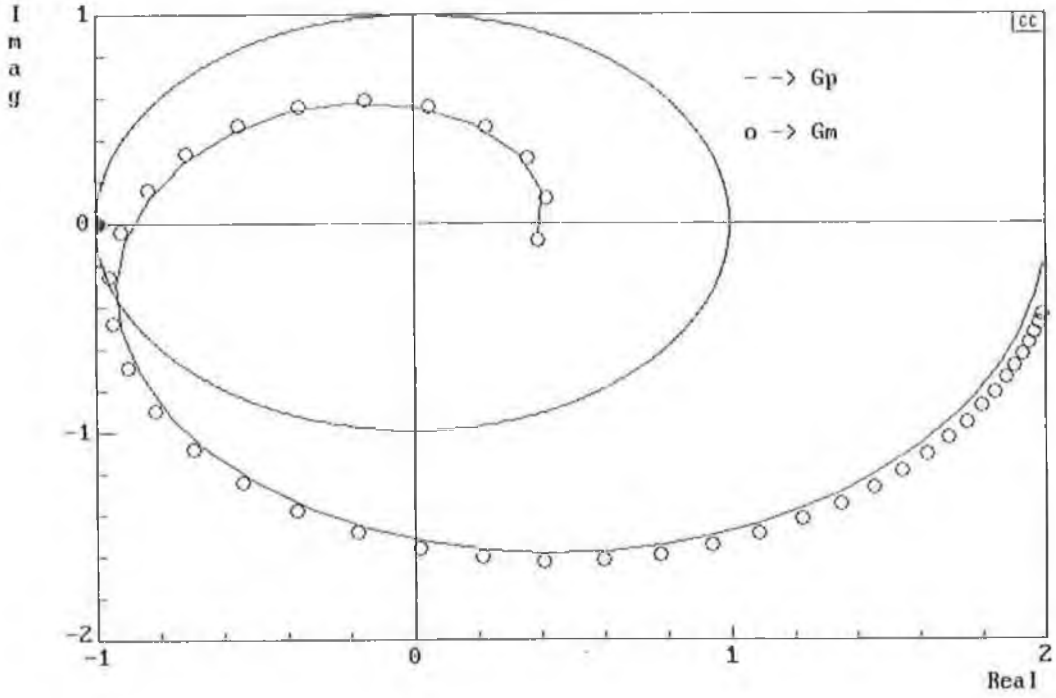


Figure 4.35: Polar plot of the process and the FOLPD model



For case B, $G_p = 2e^{-1.0s}/1 + 4.5s + 4.5s^2$. From Table 4.11, the parameters of a SOSPD model are estimated (Figures 4.36 to 4.39), following the procedure in Option 3 (Section 4.3.3.2). Fitting in the time domain and in the frequency domain, using Program CC, is shown in Figures 4.40 and 4.41, respectively.

Figure 4.36: $K_m - \mu = 0.1$

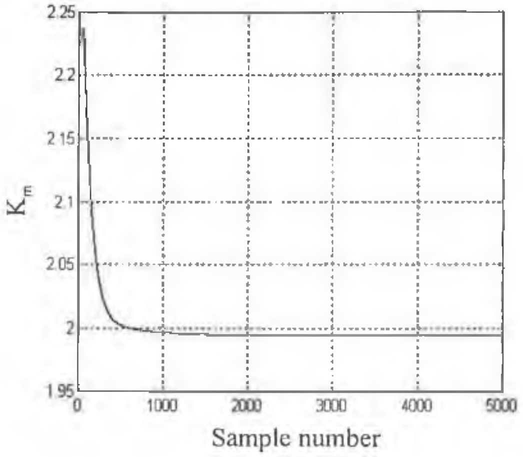


Figure 4.37: $a_{1m} - \mu = 0.1$

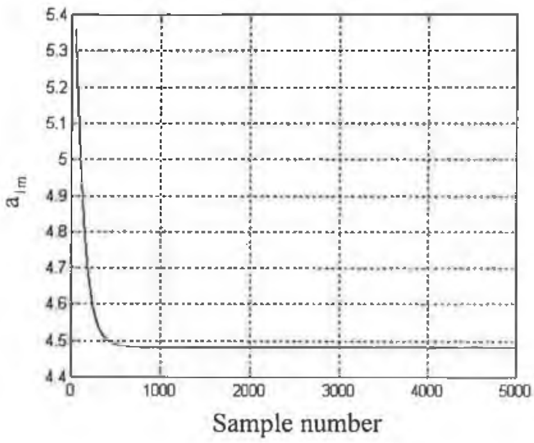


Figure 4.38: $a_{2m} - \mu = 0.1$

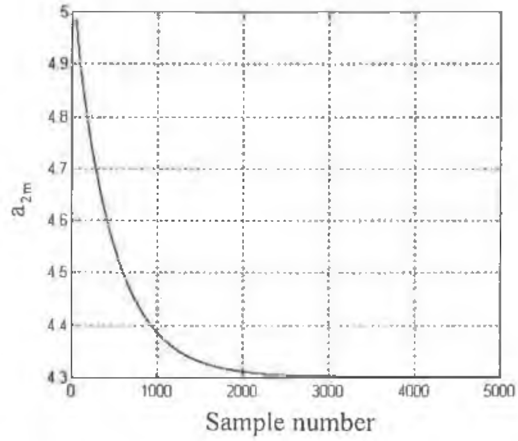


Figure 4.39: $\tau_m - \mu = 0.01$

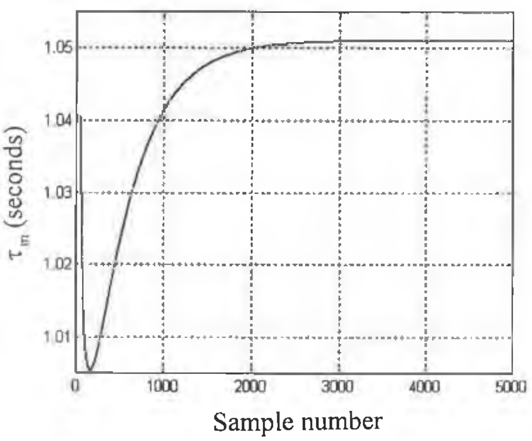


Figure 4.40: Unit step response of the process and the SOSPD model

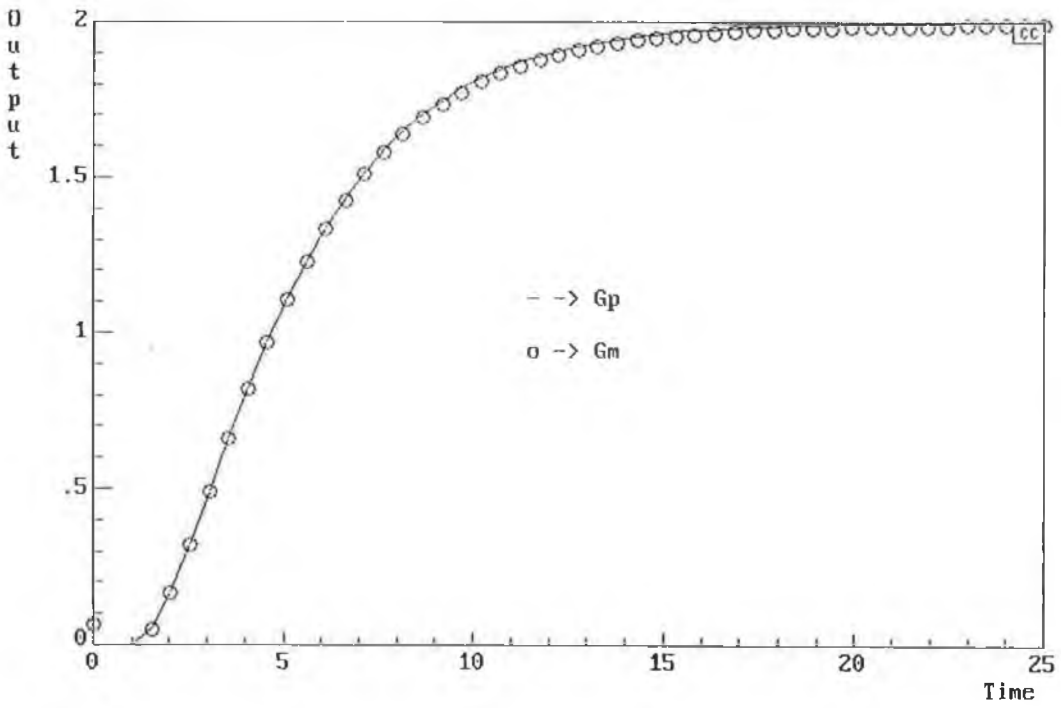
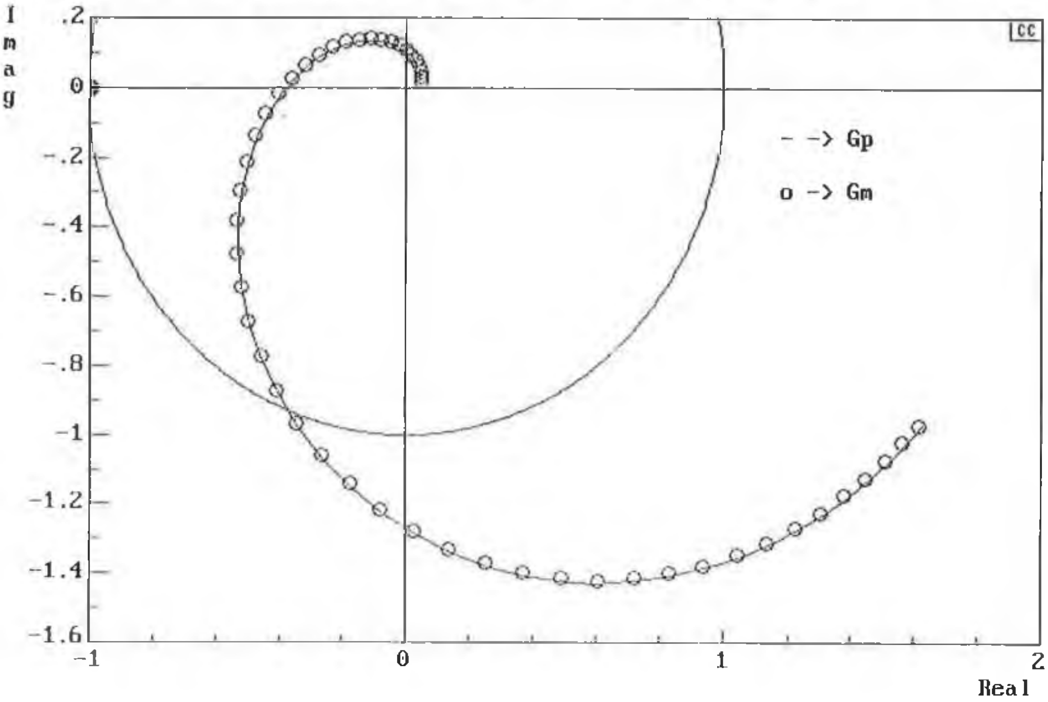


Figure 4.41: Polar plot of the process and the SOSPD model



For case E, $G_p = 2e^{-1.0s}/1+s+s^2$. From Table 4.11, the parameters of a SOSPD model are estimated (Figures 4.42 to 4.45), following the procedure in Option 3 (Section 4.3.3.2). Fitting in the time domain and in the frequency domain, using Program CC, is shown in Figures 4.46 and 4.47, respectively.

Figure 4.42: $K_m - \mu = 0.1$

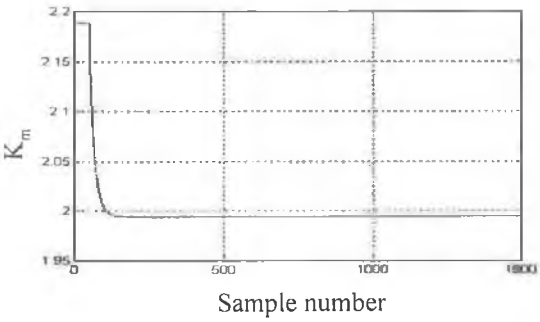


Figure 4.43: $a_{1m} - \mu = 0.1$

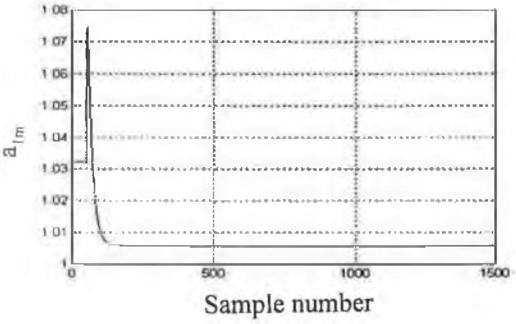


Figure 4.44: $a_{2m} - \mu = 0.1$

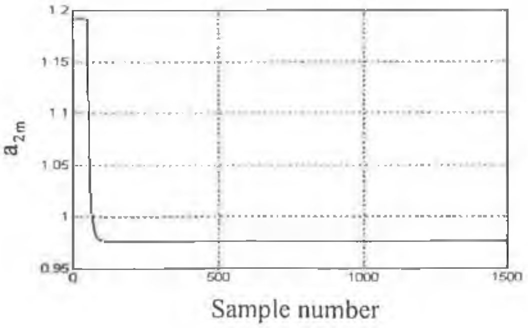


Figure 4.45: $\tau_m - \mu = 0.01$

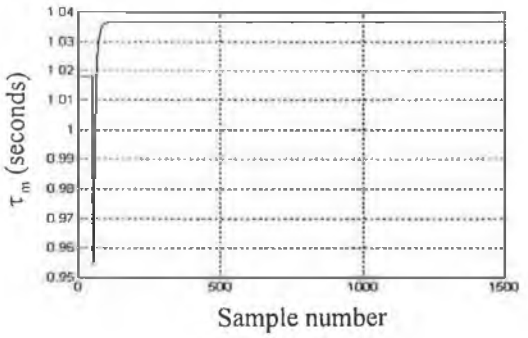


Figure 4.46: Unit step response of the process and the SOSPD model

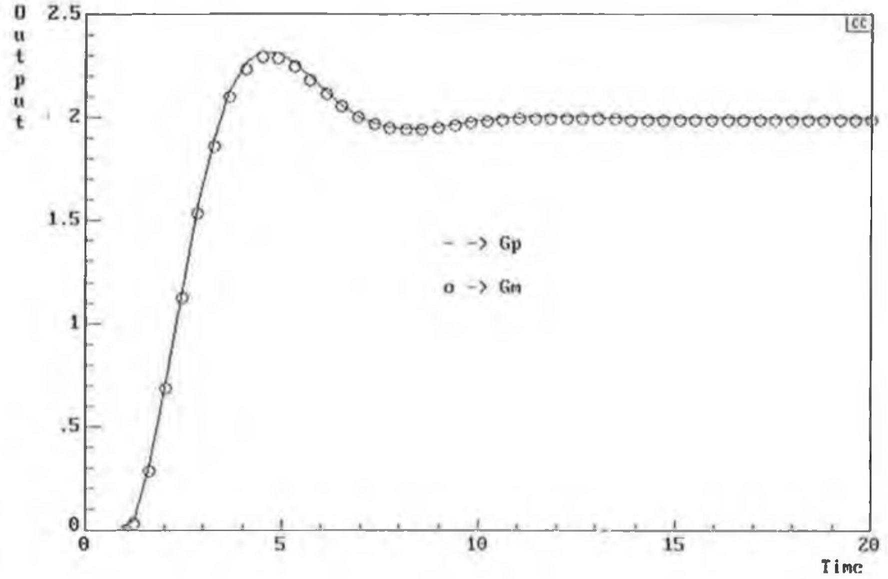
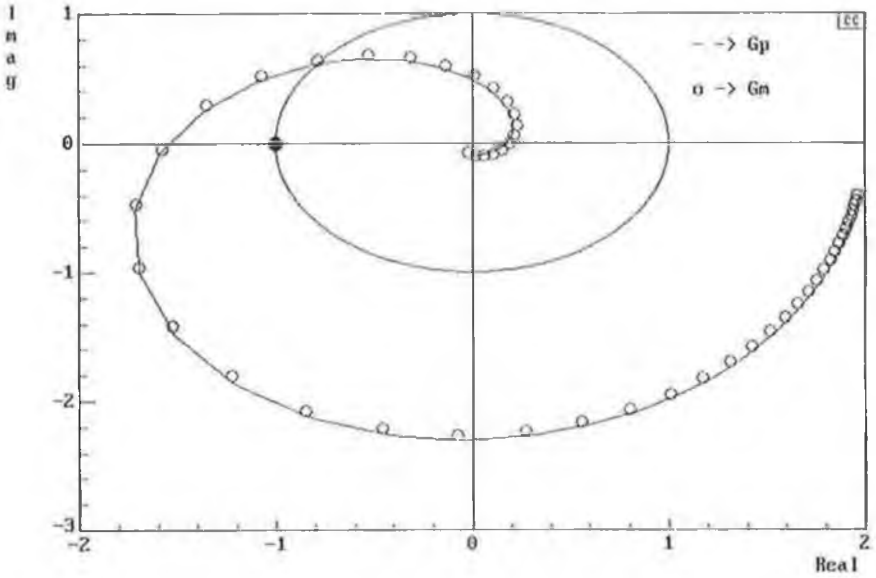


Figure 4.47: Polar plot of the process and the SOSPD model



For case F, $G_p = 2(1+2.25s)e^{-1.0s}/1+8.5s+22.5s^2+18s^3$. From Table 4.11, the parameters of a SOSPD model are estimated (Figures 4.48 to 4.51). Fitting in the time domain and in the frequency domain, using Program CC, is shown in Figures 4.52 and 4.53, respectively.

Figure 4.48: $K_m - \mu = 0.1$

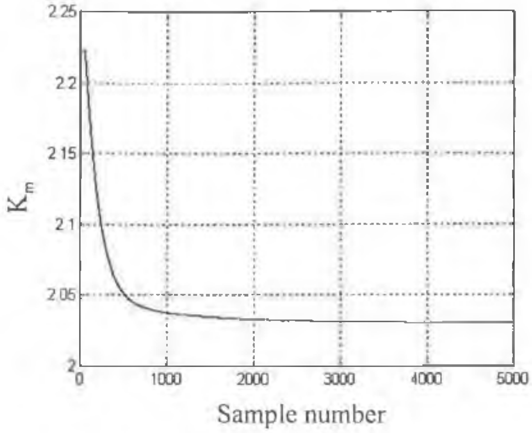


Figure 4.49: $a_{1m} - \mu = 0.1$

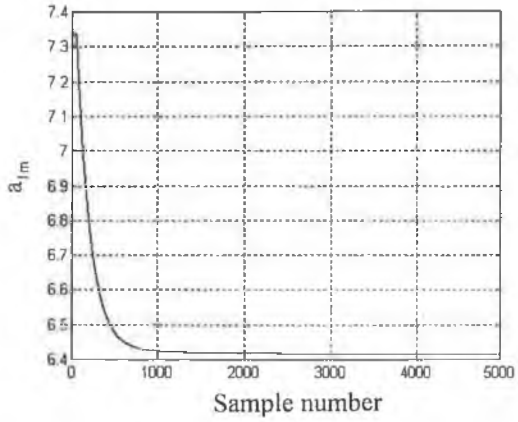


Figure 4.50: $a_{2m} - \mu = 0.1$

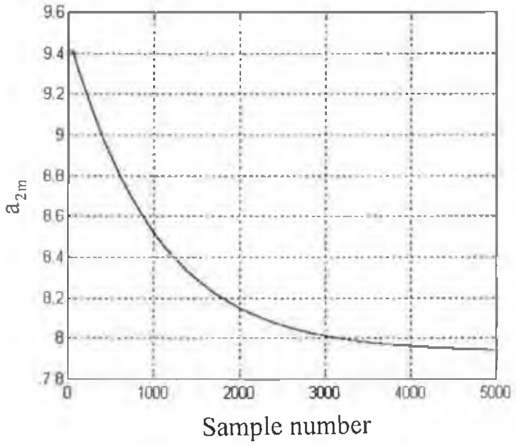


Figure 4.51: $\tau_m - \mu = 0.01$

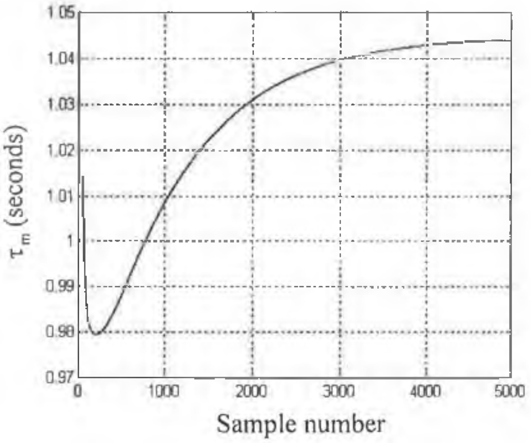


Figure 4.52: Unit step response of the process and the SOSPD model

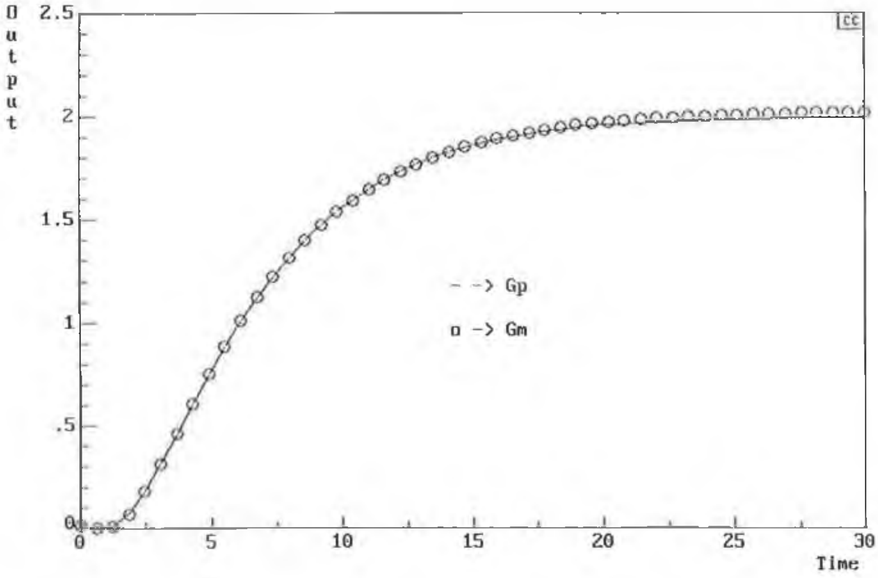
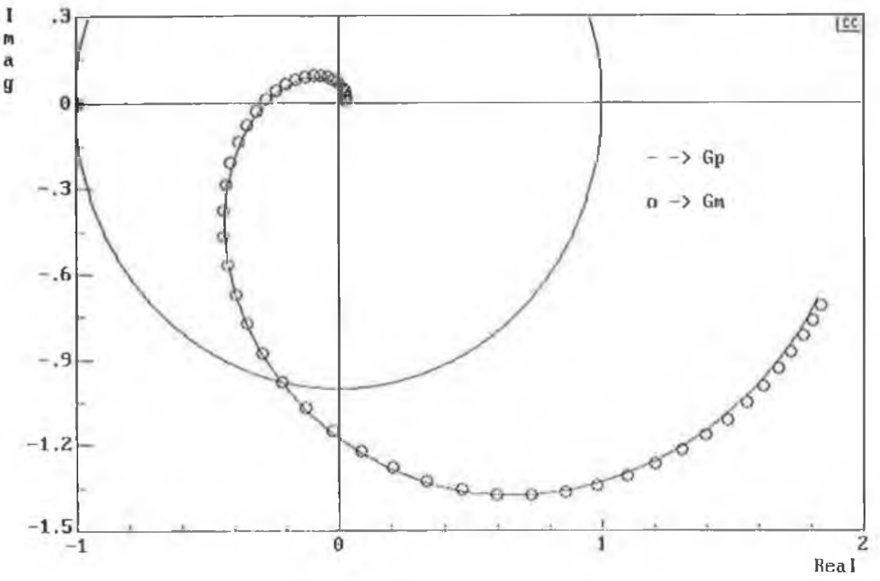


Figure 4.53: Polar plot of the process and the SOSPD model



For case G, $G_p = 2(1-2.25s)e^{-1.0s}/1+8.5s+22.5s^2+18s^3$. From Table 4.11, the parameters of a SOSPD model are estimated (Figures 4.54 to 4.57). Fitting in the time domain and in the frequency domain, using Program CC, is shown in Figures 4.58 and 4.59, respectively.

Figure 4.54: $K_m - \mu = 1.0$

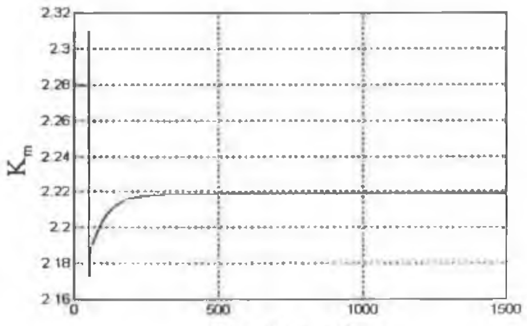


Figure 4.55: $a_{1m} - \mu = 10$

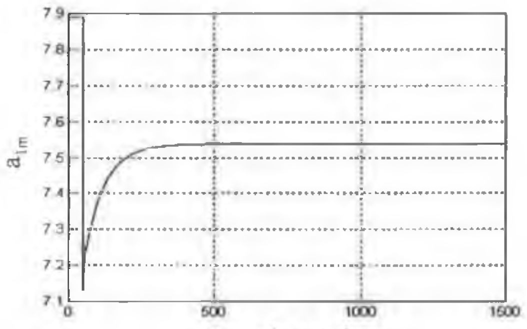


Figure 4.56: $a_{2m} - \mu = 10$

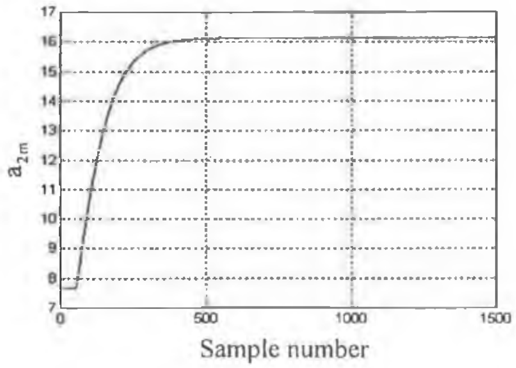


Figure 4.57: $\tau_m - \mu = 0.1$

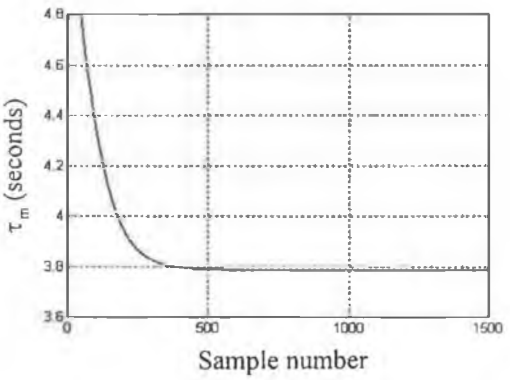


Figure 4.58: Unit step response of the process and the SOSPD model

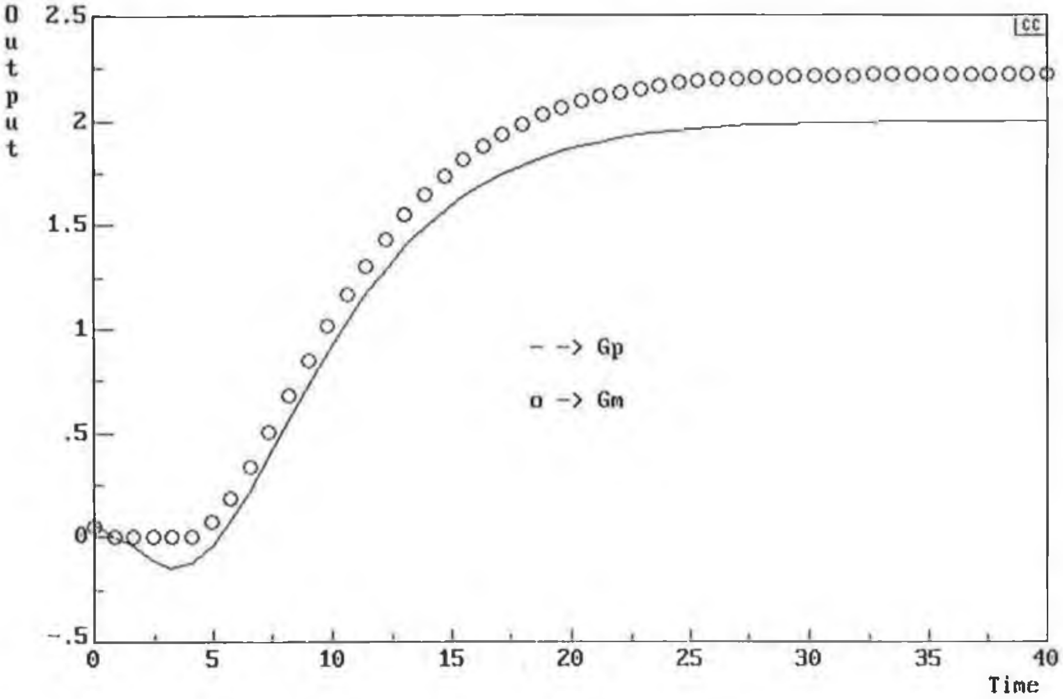
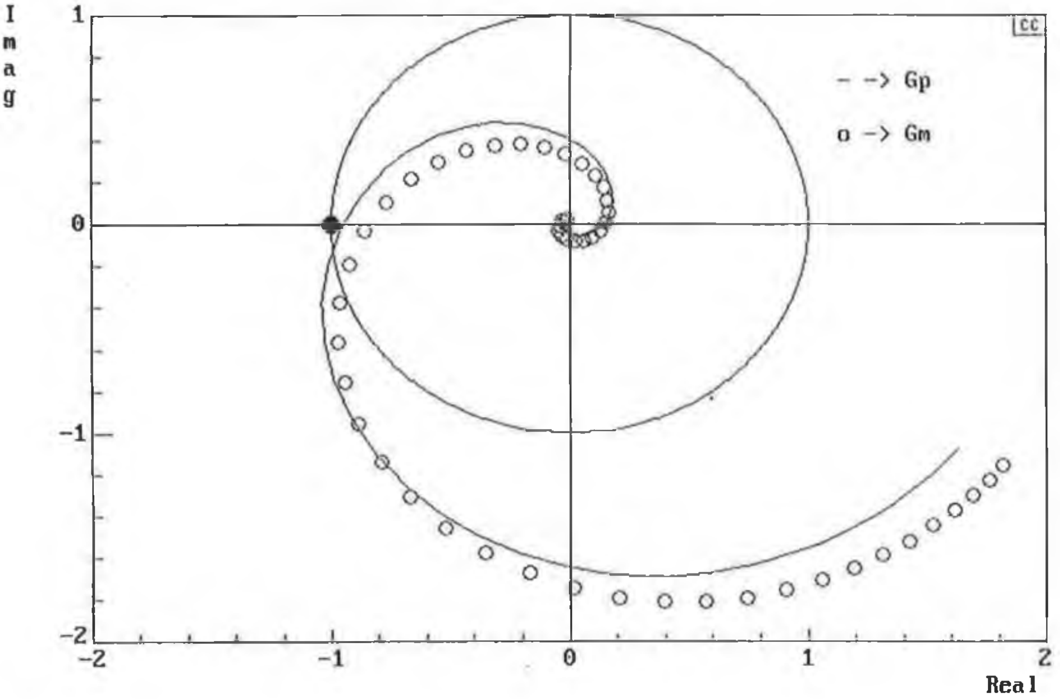


Figure 4.59: Polar plot of the process and the SOSPD model



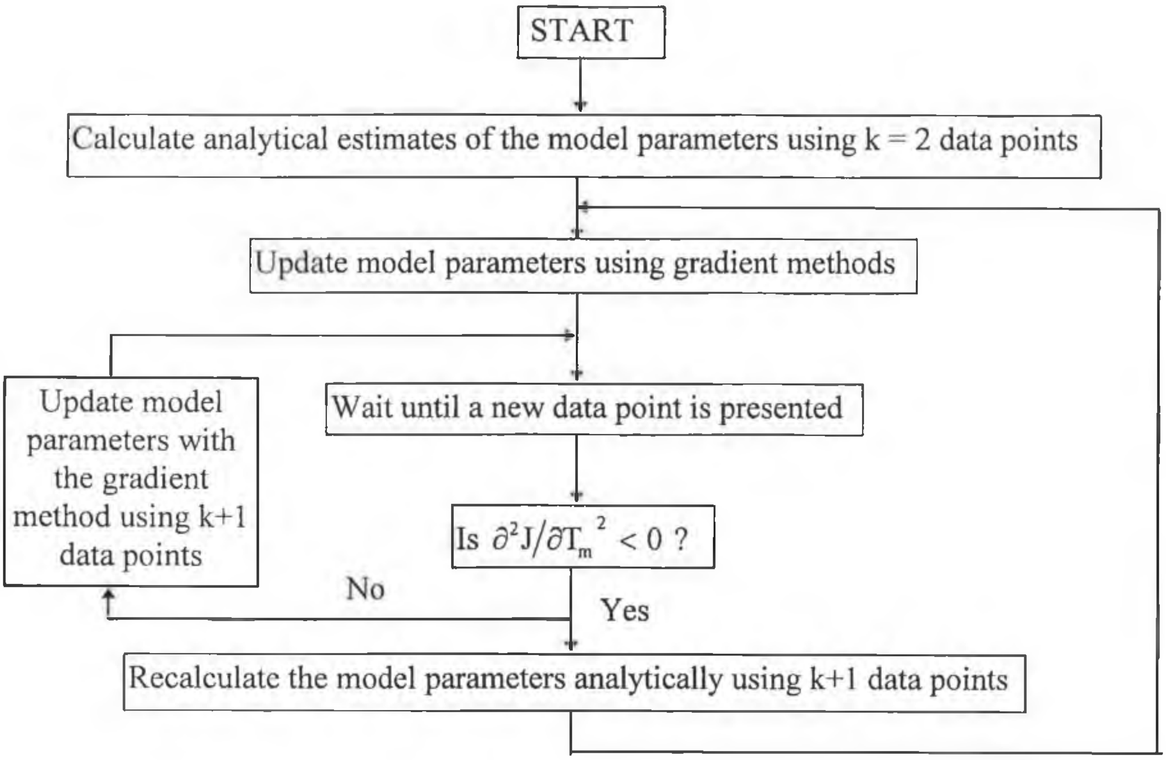
All of these results show the appropriateness of the strategy under consideration.

4.3.6 Recursive estimation of the model parameters

In the work described in the chapter thus far, it has been assumed that ten points on the process frequency response have been available for the calculation of the model parameters. It is possible that data points may become available at different times. It would be possible to wait until all the data points became available before estimating the model parameters; an alternative would be to implement a recursive scheme which would estimate appropriate model parameters as each data point became available. For example, if the parameters of a FOLPD model are to be estimated, a minimum of two data points are required to estimate the model gain analytically. However, the sensitivities of the subsequent parameter estimates to errors in the magnitude values recorded are likely to be high. One could re-estimate the analytical parameter estimates or one could rely on the gradient technique to appropriately update the parameter estimates, as more data points became available; however, the convergence of the parameter estimates to their optimum values is not guaranteed.

The algorithm proposed for the recursive estimation of the FOLPD model parameters is given in Figure 4.60.

Figure 4.60: Recursive estimation algorithm



This algorithm ensures that the condition for unimodality is fulfilled. One representative simulation result (for Case A), using the algorithm proposed in Figure 4.60, is provided in Figures 4.61 to 4.63, under the ±10% noise condition. A new data point is added every 200 sample periods. The figures show that the estimates calculated converge to the estimates determined when ten data points were used to calculate the parameters non-recursively (Figures 4.31 to 4.33). Figures 4.61 to 4.63 also show that there is a refinement in the parameter estimates as new data points are added, demonstrating the recursive nature of the method. Of course, the computational burden increases as the number of data points increase; the computational burden could be limited by defining a maximum number of data points over which the parameters should be calculated. If more data points become available, these could replace existing data points used to calculate the parameters. The extension of the method to the estimation of the parameters of a higher order model plus time delay is a natural progression of the method, though the computational burden is likely to be increased (because of the larger number of points required to estimate the parameters, combined with the greater number of second partial derivatives that would need to be compared to zero).

Figure 4.61: $K_m - \mu = 0.5$

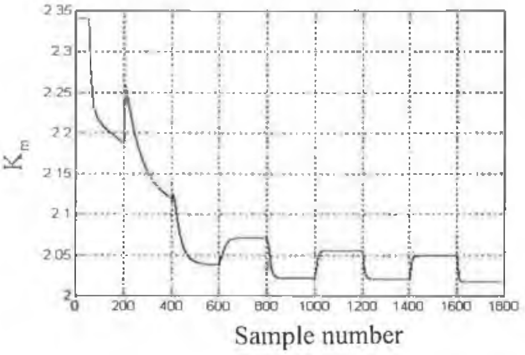


Figure 4.62: $T_m - \mu = 0.5$

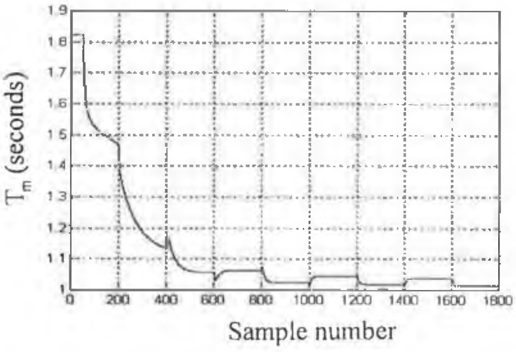
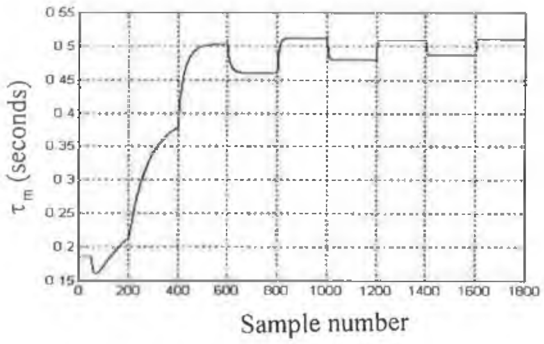


Figure 4.63: $\tau_m - \mu = 0.5$



4.3.7 Other issues

4.3.7.1 The choice of the learning rate, μ

Appropriate estimates of the learning rate, μ, (equation (4.103)) have been determined in simulation. The best setting of this value to allow rapid convergence of the parameter estimates appears to be related to the process order and to whether the process is underdamped or overdamped. Unfortunately, it is very possible for the model parameters to converge to non-optimum values if the value of the learning rate is too large. An ultimate aim would be to allow the learning rate to be adaptive. A trial and error procedure to choose the learning rate was the only satisfactory method developed.

4.3.7.2 Normalising used in the cost function

The normalising used in the cost function (equations (4.100), (4.101)) has the effect of weighting the cost function more equally over a wide range of frequencies. This facilitates the convergence of the model parameters to their optimum values, using the gradient method, over a wider range of initial model parameters than if no cost function weighting is used. Other normalising matrices, based on a different weighting of the cost function over the range of frequencies, have been employed to less effect than the normalising matrix that is used in equation (4.101). Palmor and Blau (1994) also use a normalising matrix in an effort to balance the components that make up the cost function; their matrix also involves dividing the magnitude components by the appropriate process magnitude value, though they divide the phase components by unity. The normalising used in equation (4.101) involves dividing the phase lag related components by the appropriate frequency; this is done to approximately balance out the cost function over all the phase terms.

4.3.7.3 Other methods of calculating initial model parameter values

Alternative analytical procedures, based on calculating some or all of the non-delay parameters from the phase response (instead of just calculating the time delay from the phase response) could be employed to calculate initial model parameter values for use in the gradient method. However, a disadvantage of such a scheme, particularly if the parameters of a FOLPD or SOSPD model are estimated, is the non-linear dependence of the phase on the numerator and denominator parameters of the model (equation (4.85)), which means that the solution of the simultaneous equations involved would need to be done numerically (in contrast to the present analytical determination of the parameters, when the model is in FOLPD or SOSPD form).

Alternatively, different models of the process, with corresponding analytical procedures, could be defined to calculate the model parameters; one example of such a method, which involves assuming that the model denominator has repeated poles, has already been formulated (equations (4.86) to (4.91)). Other such methods could also be implemented; for example, the model denominator $D_m(s)$ could be assumed to be as follows:

$$D_m(s) = \prod_{n=1}^{N}(1 + nT_m s) \qquad (4.164)$$

Another alternative is

$$D_m(s) = \left(1 + T_m s\right)^{N/2}\left(1 + 2T_m s\right)^{N/2} \qquad (4.165)$$

The use of equations such as (4.164) and (4.165) may change the trade-off discussed in Section 4.3.2.1, in which poorer accuracy of the parameters estimated using such procedures, compared to using the procedures to analytically determine the parameters of a general order model, is traded off against the computational intensity of the latter algorithm. It does appear that the use of the latter analytical formulae is generally indicated, provided the complexity of the equations involved to estimate the model parameters is not prohibitive.

4.3.7.4 The choice of model parameter estimation method

Tables 4.1 to 4.5, 4.7, 4.9 and 4.10 demonstrate that, in the absence of noise, the parameters determined using the analytical approach are as appropriate as the parameters determined using the two-stage approach (on balance). This is as expected, though the evaluation of the gain is problematical in the analytical approach in some cases; this is due to the relatively restricted range of process phase values over which the parameters are calculated. The advantages of the analytical approach are that it is much less computationally intensive than the two-stage approach and that there are no concerns about choosing the most appropriate value of the learning rate, μ. As expected, the analytical model parameters facilitate much poorer fitting than do the parameters calculated using the two-stage method, in both the time and frequency domains, under the ±10% noise condition. This result is compatible with the large values of the parameter sensitivity functions calculated when an analytical method is used to estimate the FOLPD model parameters (equations (4.121) to (4.124)).

4.3.8 Conclusions

1. The two-stage method defined in this section of the chapter has successfully allowed the estimation of the parameters (including the time delay) of SISO process models, from an appropriate number of arbitrarily specified points on the process frequency response, in a wide variety of simulations. Convergence of the initial model parameter estimates, calculated using the analytical approach, to the optimum model parameter estimates, calculated using the gradient approach, is possible if the initial model parameters are sufficiently close to the optimum parameters. It was not possible to prove such convergence properties analytically, though it is possible to evaluate the likelihood of convergence of the initial model parameter estimates to the optimum values by calculating the second partial derivative(s) of the cost function with respect to the denominator parameter value(s); if these expressions are greater than zero, convergence is possible, though not guaranteed. It is also possible, having obtained the analytical estimates (and determining that one or more of the second partial derivative are less than zero), to adjust them in a manner likely to allow the corresponding second partial derivatives to be greater than zero and thus to increase the possibility of convergence. An alternative simple strategy that involves the commencement of iteration at different values of the parameter estimates could also be employed to increase the probability that the parameters estimated using the gradient approach will correspond to the global minimum of the cost function.

2. In general, the required frequency range over which the process and the model must be fitted, and the acceptability of the fitting of the process to the model, depends on the

use to which the model is applied; it appears reasonable that, for many applications, fitting of the process in phase lags between 0^0 and 270^0 will be the maximum range over which good fitting is required. This is true for many compensation strategies (e.g. PID controller design); in addition, most processes, being low pass in nature, will have a small magnitude at larger phase lags, making the measurement problem greater. These considerations provide a cogent argument for estimating the parameters of either a FOLPD model or a SOSPD model; in addition, the estimation of the parameters of such a low order model will have the advantages of a lower computational burden, faster convergence and a larger parameter space for which the cost function is unimodal, compared to the estimation of the parameters of a higher order model.

3. In future work, it would be worth investigating the robustness of analytical estimates of the process parameters, calculated using data at the ultimate frequency. Frequency response data determined at the ultimate frequency is of particular interest, because if the magnitude and phase of the process is to be estimated using either the power spectral density approach or the approach that involves the ratio of the Fourier transforms of the output signal and the input signal to the process, the estimation in open loop or closed loop is unbiased by the presence of frequency terms at multiples of twice the frequency of evaluation (provided, of course, that the required conditions are observed on the excitation signals). The ultimate frequency is also in the range of frequencies over which the process model should be estimated with accuracy, for controller design purposes.

CHAPTER 5

The compensation of processes with time delay

5.1 Introduction

The compensation of processes with time delay, both in the continuous time domain and the discrete time domain, has been treated comprehensively in the published literature. Isermann (1989), (1991) classifies the two major types of compensator designs for processes with time delay, in the discrete time domain, as (a) the design of parameter optimised controllers, in which the controller structure is fixed, and the controller parameters are adapted to the controller structure and (b) the design of structurally optimised controllers, in which the controller structure and the controller parameters are adapted optimally to the structure and to the parameters of the process model. It appears reasonable that such a classification may be extended to classify compensation methods in the continuous time domain. The focus of this chapter is to review the work done in the design of both parameter optimised compensators and structurally optimised compensators, in both the continuous time and discrete time domains; it will include consideration of more recent work than that covered by Isermann (1989), (1991).

The two major classes of parameter optimised compensators will be considered, namely those based on a PID structure and those based on a lead/lag structure. The methods used to define the appropriate compensator parameters will be classified and outlined. The discussion in this section of the chapter is further detailed by O’Dwyer (1996b).

The design of structurally optimised time delay compensators will then be discussed; examples of the compensators discussed will include the Smith predictor and direct synthesis compensators. The discussion in this section of the chapter is further detailed by O’Dwyer (1996c).

5.2 Parameter optimised controllers

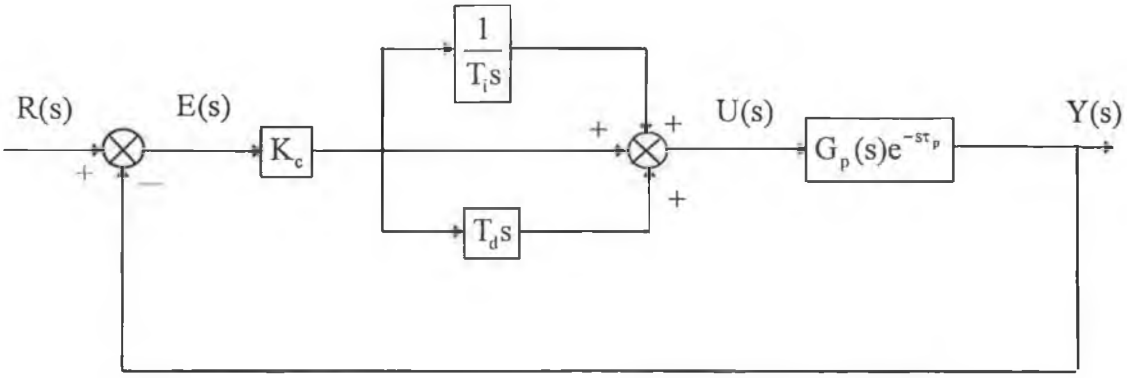
Parameter optimised controllers refer to controllers whose structure is fixed. The PID controller and the lead-lag controller are the two main types of parameter optimised controllers. The design of these controllers for processes with time delay are based on methods that were originally used for the design of the controllers for delay-free processes; in these applications, the controllers are popular, due to their relatively wide applicability and ease of use. This section of the chapter will consider how the design methods may be applied to the specification of parameter optimised controllers for processes with time delays.

5.2.1 The design of PID parameter optimised controllers

5.2.1.1 Introduction to the PID controller

The most common controller structure in process control applications is the PID (proportional integral derivative) or three term controller structure and its variations (P, PI or PD structures). The ideal continuous time domain PID controller for a SISO process model is shown in Figure 5.1.

Figure 5.1: Block diagram of a SISO process controlled by an ideal PID controller



Such a controller is expressed in the Laplace domain as follows:

$$U(s) = G_c(s)E(s) \tag{5.1}$$

with

$$G_c(s) = K_c\left(1 + \frac{1}{T_i s} + T_d s\right) \qquad (5.2)$$

and with K_c = proportional gain, T_i = integral time constant and T_d = derivative time constant. If $T_i = \infty$ and $T_d = 0$ (i.e. P control), then it is clear that the measured value, Y, will always be less than the desired value, R (for processes without an integrator term, as a positive error is necessary to keep the measured value constant, and less than the desired value). The introduction of integral action facilitates the achievement of equality between the measured value and the desired value, as a constant error produces an increasing controller output. The introduction of derivative action means that changes in the desired value may be anticipated, and thus an appropriate correction may be added prior to the actual change. Thus, in simplified terms, the PID controller allows contributions from present controller inputs, past controller inputs and future controller inputs.

It is, however, uncommon to implement the PID controller structure in equation (5.2) in practice (Astrom and Wittenmark (1984)). These authors describe more common PID controller structures, examples of which are outlined below.

1. A PID controller with a filter on the differentiator. Such a structure reduces the tendency of the differentiator to amplify high frequency noise.
2. A PID controller with derivative action on the measured value only. Such a controller does not cause a large control signal following a step change in the desired value (Kaya and Scheib (1988) label this the non-interacting controller).
3. The I-PD controller structure, in which the proportional and derivative terms act on the measured value.

The PID controller is often implemented in the discrete time domain; the ‘position’ form of the PID algorithm may be directly calculated by appropriate substitutions, using finite differences, for the integral and derivative terms. An alternative implementation is the ‘velocity’ form of the PID algorithm, in which the increment in the controller output is calculated. Two practical issues associated with the implementation of these algorithms are bumpless transfer between manual and automatic control operation, and anti-reset windup. Other issues associated with the implementation of such discrete time algorithms (such as the choice of the sample period used) are discussed in detail by O’Dwyer (1996b).

The use of the PID controllers (in any of the structures indicated) for the control of both time delayed and non-time delayed processes is very common in industrial applications. Ho *et al.* (1995a) declare that over 90% of industrial controllers are of PI type. Koivo and Tanttu (1991) suggest that there are perhaps 5-10% of control loops that cannot be controlled by SISO PID controllers (the authors suggest the use of MIMO PID controllers for these cases); in particular, the PID controller performs well if the performance requirements are modest (Hwang (1993)). PID controllers have some robustness to incorrect process model order assumptions (Lammers and Verbruggen (1985)), and limited process parameter changes. The PID controller also has the considerable advantage of being easy to understand, with tuning rules that have been validated in a wide variety of practical cases. It may also be shown that the PID controller is optimal for the control of a second order process.

However, many authorities state that PID controllers have their limitations; Deshpande and Ash (1983), for instance, suggest that PI control is used in the majority of industrial applications, as the PID controller is sensitive to process and measurement noise, is less forgiving of process parameter changes and is more difficult to tune. The PID structure and its variations are stated by these authors not to be well suited for the control of (a) non-linear processes (b) interacting processes (c) processes with the time delay greater than the time constant (for a FOLPD process model) and (d) processes with disturbances that have a significant frequency content around the systems resonant frequency of oscillation. Hagglund and Astrom (1989), (1991) also suggest that the PID controller is unsuitable for the control of (a) processes whose dynamics are dominated by a time delay (b) processes with oscillatory dynamics and (c) processes with significant stochastic disturbances; in another paper, Astrom *et al.* (1993) state that the PID controller requires unnecessarily fast sampling (if implemented in the discrete time domain). Isermann (1989) suggests that the PID controller implementation is recommended for the control of processes of "low to medium" order, with small time delays, when controller parameter setting must be done using tuning rules and when controller synthesis may be performed a number of times.

These arguments make it clear that the PID controller structure (and its variations) have a role to play in the control of some processes with time delay. In many applications, when the process is accurately modelled by a low order model plus time delay, the PID controller is recommended when the time delay is not the dominant model parameter. In subsequent sections of this chapter, the design of PID controllers for processes with time delays will be discussed in detail.

5.2.1.2 The specification of the controller parameters

Isermann (1989) outlines a number of criteria for determining the parameters of the PID controller or its variations, as follows:

(a) By using iterative methods (considered in detail in Section 5.2.1.2.1)

(b) By using tuning rules (considered in detail in Section 5.2.1.2.2)

(c) By minimising a performance criterion (considered in detail in Section 5.2.1.2.3) and

(d) By using a pole placement strategy (considered in detail in Section 5.2.1.2.4).

Isermann (1989) states that the design methods described in (a) and (b) are suitable if there is no specification for the control performance, or if the process has simple behaviour or a low settling time. Isermann (1989) also suggests that the design methods described in (c) and (d) are suitable for controller design where there are "stringent" performance requirements or slow, complex or changing process behaviour. However, the applicability of the PID controller is questionable in such implementations, particularly if the process has a significant time delay.

The control of processes of variable or uncertain structure or parameters may be handled in two ways:

(1) A self-tuning controller may be designed to allow appropriate variation in the controller parameters as the process parameters vary; explicit self-tuning controllers are normally designed if the variations in the process parameters are known, and implicit self-tuning controllers are normally designed otherwise. This discussion does not consider self-tuning PID controllers separately from fixed parameter PID controllers, as the model based, analytical design techniques for both (discussed in Sections 5.2.1.2.3 and 5.2.1.2.4) tend to be the same.

(2) The controller may be designed to be robust to process parameter variations. There is a certain amount of robustness associated with the PID controller implementation, which may be assessed after an appropriate controller has been designed. Alternatively, the controller may be designed to satisfy simple robustness measures (e.g. gain margin or phase margin specifications); such specifications are considered in the PID controller designs discussed in Sections 5.2.1.2.1 to 5.2.1.2.4, inclusive. More

generally, a more complete robust design strategy based on, for example, an Internal Model Control (IMC) procedure may be employed; these strategies are considered in Section 5.2.1.2.5.

5.2.1.2.1 Iterative methods

The choice of appropriate compensator parameters may be achieved either experimentally (for example, by manual tuning), or by using a graphical approach in either the time domain (typically using root loci) or the frequency domain (typically using Bode plots). Perhaps a majority of tuning methods could be considered iterative, as the effect of the compensator designed is often assessed, and changes made to the compensator as appropriate; this section of the chapter will consider only those methods in which a number of trials are typically needed to achieve a satisfactory compensator.

A typical experimental approach to trial and error tuning is discussed by Seborg *et al.* (1989), who suggest the following approach:

(a) Set T_d to a minimum value and T_i to a maximum value.

(b) Set K_c at a low value and put the controller on automatic (if appropriate).

(c) Increase K_c by small increments until continuous cycling occurs after a small set point change or load change. Set K_c to half this value.

(d) Decrease T_i by small increments until continuous cycling re-occurs. Set T_i to three times this value.

(e) Increase T_d until continuous cycling re-occurs. Set T_d to one third of this value.

Other experimental approaches are outlined by Pollard (1971), Power and Simpson (1978), Deshpande and Ash (1983), Leigh (1987) and De Santis (1994). However, trial and error tuning using an experimental approach has some significant disadvantages; Seborg *et al.* (1989) list these as

(a) The time consuming nature of the method and

(b) The requirement that the process has to be driven to its stability limit.

In addition, the authors declare that the strategy is not applicable to the design of compensators for processes that are open loop unstable.

A graphical trial and error approach to controller tuning may be done in either the time domain or the frequency domain. The time domain design is done using root locus diagrams. In this method, the process is assumed to be well modelled by a second order model; thus, it is the location of the process dominant poles that are considered to be significant. Such root locus based approaches may be used to design appropriate PD controllers (Kuo (1991), Philips and Harbor (1991)), PI controllers (Dorf (1989), Kuo (1991)) and PID controllers (Kuo (1991), Philips and Harbor (1991)); it is, however, questionable that a process with time delay would be sufficiently well modelled by a second order model to allow the design of appropriate compensators. The frequency domain design is typically done using Bode plots; unlike the time domain procedure, the process is not necessarily modelled by a second order model. The frequency domain may also be used to design PD controllers (Kuo (1991)), PI and PID controllers (Kuo (1991), Philips and Harbor (1991)).

Alternatively, an analytical approach to the design of PD, PI or PID controllers may be employed. In the time domain, the design criterion is typically that a root location should be on the root locus, with the frequency domain design criterion typically being the achievement of a desired phase margin (Philips and Harbor (1991), Shahian and Hassul (1993)). The latter authors show, in one simulation result, a step response of the compensated system in servo mode that has an overshoot 22% larger than the specification, indicating that the method used (in the time domain) must be considered iterative; such inaccuracies are typical, because of inaccurate process models used. Rao and Perdikaris (1988) also discuss these techniques. Similar methods may also be described in the discrete time domain (Shahian and Hassul (1993)), though Seborg *et al.* (1989) declare that the discrete root locus method is not recommended for designing PID controllers, as the dynamic response of sampled data systems is not uniquely determined by the closed loop pole locations.

In conclusion, the iterative methods for controller design discussed provide a first approximation to desirable controller parameters. The methods described are perhaps most suitable if there are simple specifications on the performance of the closed loop control system to be achieved; they may be applied to the control of processes with time delays, though the accuracy of the process model used in the controller design is an issue.

5.2.1.2.2 Tuning rules

Many authors have defined tuning rules for the specification of the PID controller parameters. In most cases, the motivations for using tuning rules are the ease of use of the rules and the difficulty in calculating a sufficiently accurate process model to employ other, more accurate, tuning strategies. Therefore, tuning rules provide non-optimum controller parameters in most cases. This section of the chapter considers (a) tuning rules based on the measured step function (also called process reaction curve methods) and (b) tuning rules based on recording appropriate parameters at the ultimate frequency (also called continuous cycling methods). Tuning rules based on minimising a performance criterion are considered in Section 5.2.1.2.3.

Process reaction curve methods for controller design are based on calculating the controller parameters from the model parameters that are determined from the step response of the process in open loop. This method was originally suggested by Ziegler and Nichols (1943), who model the process by a FOLPD model, estimate the model parameters using a tangent and point method (as indicated in Chapter 2) and define appropriate tuning parameters for the P, PI and PID control of the process. This method achieves the quarter decay ratio criterion (approximately); this criterion refers to the maintenance of the ratio of the amplitudes of two successive oscillations (of the closed loop output) at 0.25, for both set point and disturbance input changes. Shaw (1993) comments that the rules facilitate good compromise between instability and sluggish control; Hang (1989) suggests that, for PID regulator design, the Ziegler-Nichols tuning strategy positions the PID controller zeroes relative to the process poles (in many cases) to achieve approximately optimal disturbance rejection response. Other tuning rules of this type are defined by Cohen and Coon (1953), which also facilitate an approximate quarter decay ratio response in servo and regulator modes, though the response tends to be more aggressive than that achieved with the tuning rules of Ziegler and Nichols (1943); both sets of tuning rules apply for values of the time delay in a range of 0.1 to 1.0 times the time constant (Smith and Corripio (1985)). Yuwana and Seborg (1982), Deshpande and Ash (1983), Smith and Corripio (1985), Hwang and Chang (1987), Ringwood (1987), Seborg *et al.* (1989), De Paor (1993), Shaw (1993) and Ho *et al.* (1995a) discuss the characteristics and operation of the tuning rules of Ziegler and Nichols (1943) and/or Cohen and Coon (1953), in detail. Seborg *et al.* (1989) suggest that the advantages of tuning strategies based on such process reaction curve methods are that only a single experimental test is necessary, a trial and error procedure is not required and the controller settings are easily calculated. However, disadvantages of the method include the difficulty of calculating an accurate and parsimonious process model, and the possibility of significant load changes occurring during the test that may distort the test results. Process reaction curve methods may also be used to tune discrete time compensators (Isermann (1989), Su (1993)) and compensators for MIMO processes (Jussila and Koivo (1987)).

Continuous cycling based tuning rules are calculated from the parameters (controller gain and oscillation period) recorded at the ultimate frequency (i.e. the frequency at which marginal stability of the closed loop control system occurs). One of the first ultimate cycle tuning methods was defined by Ziegler and Nichols (1942) (henceforth referred to as the Ziegler-Nichols ultimate cycle method) for the tuning of P, PI and PID controller parameters of a process that may or may not include a time delay. The tuning rule involves bringing the system to marginal stability under proportional control, recording the proportional gain at marginal stability and the period of oscillation of the output and calculating appropriate tuning parameters based on these values. The tuning parameters approximately facilitate the achievement of the quarter decay ratio criterion in the response to both set point inputs and disturbance inputs, for many processes. De Paor (1993) shows that the tuning rules implicitly build an adequate frequency domain stability margin into the compensated system; Krishnaswamy *et al.* (1987) declare that the rules display robustness against signal noise and process parameter variations. The main advantage of the ultimate cycle tuning strategy is that the controller settings are easily calculated.

The weakness of the method is that the system must be destabilised under proportional control, and the empirical nature of the method means that uniform performance is not achieved in general (Hwang and Tseng (1994)). Pessen (1994) states that the drawbacks of using the ultimate cycle method in controller design is that (i) several trials must typically be made to determine the ultimate gain (ii) the resulting process upsets may be detrimental to process quality, especially if the disturbances pass through to other production units in the plant and (iii) there is a danger of misinterpreting a limit cycle as representing the stability limit.

Other authors that discuss the specification and implementation of continuous cycling tuning algorithms include Harriott (1964), Pollard (1971), Weber and Bhalodia (1979), Yuwana and Seborg (1982), Astrom and Hagglund (1984), Tan and Weber

(1985), Eitelberg (1987), Hippe *et al.* (1987), Hang and Astrom (1988), Bogere and Ozgen (1989), Lee (1989), Mantz and Tacconi (1989), (1990), Astrom (1991), Hang *et al.* (1991), Hang and Sin (1991), Litt (1991), Astrom *et al.* (1992), De Carvalho (1993), Lee *et al.* (1993), Hang *et al.* (1994a), Ho *et al.* (1995a), Kim (1995) and Wang *et al.* (1995a). Continuous cycling methods may also be used to tune discrete time compensators (Kofahl and Isermann (1985), Isermann (1989), Seborg *et al.* (1989), Shahian and Hassul (1993), Ringwood and O'Dwyer (1994a), (1994b), Bobal (1995)) and to tune compensators for MIMO processes (Luyben (1986), Isermann (1991), Loh and Vasnani (1992), Loh *et al.* (1993), Shen and Yu (1994), Wu *et al.* (1994), Zhuang and Atherton (1994) and Palmor *et al.* (1995)).

Other types of tuning rules have also been defined for PID implementation; Da Silva *et al.* (1988), for instance, present a rule based self-tuning procedure based on a pattern recognition approach, while Zhao *et al.* (1993) use fuzzy rules and reasoning to determine appropriate PID controller parameters, in an on-line environment. An expert system approach to choosing the controller parameters is outlined by Li (1994). Other tuning rules for PID controllers for MIMO processes are outlined by Lieslehto *et al.* (1991).

In conclusion, the motivations for using tuning rules are the ease of use of the rules and the difficulty in calculating a sufficiently accurate process model to employ other, more accurate, tuning strategies. The design methods in this section are thus most suitable if there is a simple specification on the performance of the closed loop control system to be achieved (such as an approximate quarter decay ratio criterion), when the process has a non-dominant time delay term.

5.2.1.2.3 The minimisation of a performance criterion

A number of performance (or optimisation) criteria may be used to design appropriate PI or PID controller parameters for a process. Simple optimisation criteria include the integral of the absolute value of the error (IAE) criterion and the integral of time multiplied by the absolute value of the error (ITAE) criterion. Such criteria suffer from the disadvantage that engineering constraints such as slew rate and control effort are not explicitly taken into account in the design of the controller parameters (Hemerly (1991)). An example of a more complete optimisation criterion is as follows:

$$J = \int_0^\infty e^2(t) + \rho u^2(t)dt \tag{5.3}$$

with e(t) = error signal, u(t) = control effort, ρ = control weight. Such a quadratic performance criterion, however, has the disadvantage that it is not possible to determine, *a priori*, the relationship between the control weight and desired transient response criteria, for instance (Hemerly (1991)).

One advantage of the use of optimisation criteria is that a unique set of controller parameter values may be calculated (unlike the fulfilment of the approximate quarter decay ratio criterion, for instance). Tuning rules to minimise a performance criterion have been defined by many authors; Rovira *et al.* (1969), for instance, model the process by a FOLPD model and define tuning parameters for PI and PID control of the process, based on minimising the IAE and ITAE criteria. Optimum controller parameters are separately defined for set point tracking applications and disturbance input rejection applications. The tuning rules apply for a range of time delay values between 0.1 and 1.0 times the time constant. Tuning rules of this type are also discussed by Ohta *et al.* (1979), Jutan and Rodriguez (1984), Nishikawa *et al.* (1984), Cheng and Hung (1985), Kaya and Scheib (1988), Jutan (1989), Seborg *et al.* (1989), Shaw (1993), Wang *et al.* (1993a), (1995a), Zhuang and Atherton (1993), (1994a), (1994b), Ho *et al.* (1995), Huang and Lin (1995) and Hwang and Fang (1995). Discrete time PID compensators may also be tuned by minimising performance criteria, using tuning rules (Moore *et al.* (1969), Rovira *et al.* (1969), (1970), Huang and Chao (1982), Astrom and Wittenmark (1984), King (1984b), Isermann (1989) and Su (1993)).

The performance criteria may also be minimised analytically, or otherwise, to calculate appropriate controller parameters. Harris and Mellichamp (1985), for instance, outline a methodology to tune a PI or PID controller to met multiple closed loop criteria. These criteria are subsumed into a single performance index that depends on frequency domain parameters with the exact design of the performance index being arbitrary; the authors choose their index as a function of the resonant peak ratio, the phase margin and the maximum resonant frequency. The method reflects the important point that there is no one set of tuning values that provide the optimum response in all respects. Other analytical methods to calculate PID controller parameters include those defined by Dumont *et al.* (1985), Huang *et al.* (1985), Patwardhan *et al.* (1987), Penner (1988), Zevros *et al.* (1988), Gorecki *et al.* (1989), Hill and Venable (1989), Lee *et al.* (1990), Ruano *et al.* (1992), Hassell and Harper (1994), Schei (1994) and Abbas and Sawyer (1995). Pemberton (1972a), (1972b), in interesting papers, argues that if the time delay may be approximated by a zero, then the PI and PID algorithms may be directly derived as the optimal controller structure for a FOLPD process model and a second order lag plus time delay model, respectively. This approximation is valid over a certain frequency range; the authors specify a number of methods of determining the tuning parameters for both controllers, though the corresponding closed loop responses appear quite oscillatory in many cases (presumably because of the mismatch between the process and the model).

Discrete time PID compensators may also be tuned using optimisation criteria; Radke and Isermann (1984), for instance, define a method for calculating the PID controller parameters that involves minimising a performance criterion that is defined as the sum of the squares of the error plus weighted control effort over all samples. Other algorithms for specifying discrete time PID compensators are defined by Shigemasa and Akizuki (1981), Cameron and Seborg (1983), King (1984a), (1984b), Ralston *et al.* (1985), Isermann (1989), (1991), Katende and Jutan (1993), Yamamoto *et al.* (1994) and Poulin *et al.* (1996). Appropriate PID compensators for MIMO processes may also be defined in the continuous time domain (Zhuang and Atherton (1994a), Wang and Wu (1995), Puleston and Mantz (1995)) and in the discrete time domain (Isermann (1991)).

In conclusion, the design methods based on minimisation of a performance criterion may be divided into tuning rule based methods and analytical, or other, methods. As mentioned previously, the tuning rule based methods are suitable for controller design where there are simple performance requirements to be implemented. The other methods are suitable for controller design where there are more complex performance requirements to be achieved (Isermann (1989)), for the control of non-dominant time delay processes.

5.2.1.2.4 Direct synthesis

A direct synthesis algorithm is one that results in a controller that gives a specified closed loop response. This may be done by specifying the desired poles of the closed loop response, for instance, though more generally, the desired closed loop transfer function may be specified. Direct synthesis methods include pole placement strategies for designing appropriate PID controllers, as well as controller design techniques that allow the achievement of a specified gain margin and phase margin.

Two term or three term controller parameters may be defined in the continuous time domain, using a pole assignment strategy, in either the time domain or the frequency domain. One time domain design method is described by Chiu *et al.* (1973b), who define algorithms for the design of PI and PID controller parameters based on the desired time constant of the closed loop system response. A FOLPD process model is used if the PI controller parameters are required, and a SOSPD model is used if the PID controller parameters are required. Other time domain strategies of this type are also discussed in detail by Pemberton (1972a), (1972b), Borg and Giles (1975), Tachibana (1984), Smith and Corripio (1985), Arzen (1987), Hwang and Chang (1987), Gorecki *et al.* (1989), Schuster (1989), Seborg *et al.* (1989), Brambilla *et al.* (1990), Aguirre (1992), Hwang (1993), McAnany (1993), Ho *et al.* (1994), Hwang and Tseng (1994), Hwang and Shiu (1994), Jin (1994), Shafiei and Shenton (1994) and Jacob and Chidambaram (1996). Frequency domain design methods are described by Edgar *et al.* (1981), Sanathanan and Quinn (1987) and Barnes *et al.* (1993); the latter authors, for instance, design a PID controller for a process with time delay by minimising the sum of the squared errors between the desired and actual polar plots. The pole assignment strategy may also be used in the discrete time domain, to determine appropriate controller parameters; Chiu *et al.* (1973b), for example, define algorithms for PI and PID controller parameter selection, based on knowledge of the desired time constant (and time delay) of the closed loop system response in the discrete time domain, in an analogue of the techniques defined by the same authors in the continuous time domain. Other discrete time techniques for controller design based on pole assignment for SISO applications are discussed by Radke and Isermann (1984), Ortega and Kelly (1984), Tjokro and Shah (1985), Teng and Sirisena (1988), Keviczky and Banyasz (1988), (1992), Vermeer *et al.* (1988), Habib and Sungoor (1989), Isermann (1989), Teng (1990), Hemerly (1991), Wellstead and Zarrop (1991), Pal *et* *al.* (1992), Brown *et al.* (1993), Yang (1993) and Leva *et al.* (1994). Keviczky and Banyasz (1988), for instance, model the process as a SOSPD model in the discrete time domain and let the numerator of the PID controller equal the denominator of the process model; the remaining controller parameters are chosen to provide a compensated system phase margin of 60 degrees. Discrete time techniques for controller design based on pole assignment for MIMO applications are discussed by Koivo and Sorvari (1985), Jones and Porter (1985), Gawthrop and Nomikos (1990) and Pal *et al.* (1992).

The controller parameters may also be defined using a gain and/or a phase margin specification, in the SISO continuous time domain; Ho *et al.* (1995b), for example, derive accurate and simple analytical formulae to tune PI and PID controller parameters for FOLPD and SOSPD models, to meet gain margin and phase margin specifications. Other authors that design PI or PID controllers based on this specification include Astrom and Hagglund (1984), Tan and Weber (1985), Balchen and Lie (1987), De Paor and O'Malley (1989), Thomson *et al.* (1989), Hagglund and Astrom (1989), (1991), Landau and Voda (1992), Schei (1992), Leva (1993), Venkatashankar and Chidambaram (1994), Voda *et al.* (1994), Voda and Landau (1995a), (1995b) and Kavdia and Chidambaram (1996). A continuous time technique of controller design based on gain and phase margin specifications for a 2x2 MIMO process is discussed by Zhuang and Atherton (1994a).

In conclusion, a direct synthesis algorithm is one that results in a controller that gives a specified closed loop response. There is, of course, some overlap between the methods discussed in this section and the methods that involve minimisation of an appropriate performance criterion (e.g. the method defined by Barnes *et al.* (1993) above); both types of method are suitable for controller design where there are well-defined performance requirements to be achieved (Isermann (1989)), for the control of non-dominant time delay processes.

5.2.1.2.5 Robust controllers

The chapter thus far has not considered formally the design of PI and PID controllers in the presence of unmodelled process dynamics, or when the process parameter values drift from the model values. In some cases, the robustness of the control system to variations in these parameter values has been considered as a means of evaluating the controller strategy. Typically, the robustness of the design is checked by calculating the sensitivity of the closed loop transfer function to changes in the process parameter values as frequency changes (mapped on to a Bode plot) or by calculating the sensitivity of the pole locations to changes in the process parameter values. Such calculations are relevant for small changes in the process parameter values only (Dorf (1989), Kuo (1991)). The disadvantage of such an approach is that the design of a controller with the required robustness may be iterative. In other cases, simple robustness measures have been built into the controller design; a typical example is the design of the controllers based on the achievement of a specified gain margin and/or phase margin in the frequency domain. Other robust controller design strategies in the frequency domain are discussed by Rivera *et al.* (1986), Dorf (1989), Morari and Zafiriou (1989) and Seborg *et al.* (1989).

This section of the chapter deals more fully with the design of PID controllers, with an explicit robust stability and robust performance criterion built in to the design process. One method of designing robustness into the controller specified is to use the Internal Model Control (IMC) design procedure. Seborg *et al.* (1989) discuss this strategy, in which a model is defined for the process; the PID controller is designed by resolving the process model as follows:

$$G_m = G_m^+ G_m^- \tag{5.4}$$

with G_m^+ containing the time delay terms and all right half plane zeroes. The controller is then designed as follows:

$$G_c = \frac{f}{G_m^-} \tag{5.5}$$

with f = desired closed loop transfer function. Morari and Zafiriou (1989) show how the full IMC design procedure, which allows uncertainty on the process transfer function parameters to be specified, may be used to design an appropriate PID controller for a FOLPD process both with time delay uncertainty and with general parameter uncertainty. The controller is designed by minimising an integral of squared error (ISE) performance criterion. Robust stability and robust performance criteria are defined explicitly during the analysis. The IMC design procedure and its application to PI and PID controller design for SISO processes with time delay is also discussed by Rivera *et al.* (1986), Chien (1988), Seborg *et al.* (1989), Chien and Fruehauf (1990), Rivera and Morari (1990), Fisher (1991), Chia and Lefkowitz (1992), Tyreus and Luyben (1992), Fruehauf *et al.* (1993), Hang *et al.* (1993), Lee and Sung (1993), Peebles *et al.* (1994) and Ho *et al.* (1995b). Friman and Waller (1994) apply the IMC procedure to the design of appropriate PI and PID compensator parameters of a 2x2 MIMO process modelled by integrator plus time delay terms or by gain plus time delay terms.

Alternatively, other robust strategies may be used to design appropriate controllers for SISO processes. Kawabe and Katayama (1994), for instance, consider the problem of acceptable control performance, and closed loop stability, in the presence of process parameter uncertainties. An I-PD controller is recommended to compensate a process with time delay, and the controller parameters are adjusted to minimise the ISE criterion maximised by process parameters belonging to a bounded set. The time delay is approximated by a first order Pade approximation during the development (which, of course, adds to the uncertainty in the dynamics). Other robust methods are described by Devanathan (1991) and Al-Saggaf (1994). In an interesting recent paper, Hayes and Holohan (1996) use results from L_1 robust control theory to tune PID controllers when the plant is poorly modelled, is non-linear, has variable time delay and/or has many operating points. The approach considers both robust stability and robust performance requirements, and the controller parameters that yield the defined specification of these requirements may be estimated numerically.

In conclusion, the robustness of a particular PID controller design to process parameter variations may be assessed after the controller is designed using a variety of methods. Alternatively, both a robust stability criterion and a robust performance criterion may be built into the PID controller design, as part of the specification to be fulfilled.

5.2.2 The design of lead, lag or lead-lag parameter optimised controllers

Lead, lag or lead-lag controllers are simple alternatives to the PID controllers described in the previous sections. The designs of such controllers are described in detail by O'Dwyer (1996b); the designs are primarily based on a root locus procedure in the time domain, or using Bode plots in the frequency domain. Thus, in the time domain, the process with delay has to be modelled by a model without delay; controller design procedures for delay free systems may then be applied. Of course, a large mismatch may be present between the process and the model, and this may affect the acceptability of the controller design in an implementation environment. In the frequency domain, the time delay appears as extra phase lag; the design of the compensators for such processes is thus the same as the design of compensators for delay free processes. Little work has been done to design such compensators specifically for processes with time delays, unlike the corresponding work that has been done for the design of PID compensators for these applications.

5.2.3 Conclusions

The choice of a parameter optimised control scheme (typically a PID controller) as opposed to a more flexible control scheme depends on a number of factors, which have been outlined in Section 5.2.1. One factor that is consistently mentioned (by Deshpande and Ash (1983) and Hagglund and Astrom (1989), (1991), amongst others) is that the PID controller and its variations are not suitable for the control of a process whose dynamics are dominated by a time delay term. A common rule of thumb quoted is that PI and PID controllers are suitable for the control of a FOLPD process if $0.1 \le \tau_m / T_m \le 1.0$. O'Dwyer (1996b) discusses this topic in detail, quoting authors that suggest more detailed rules of thumb, such as Astrom (1991) and Astrom *et al.* (1992).

In an interesting perspective on this issue, Shinskey (1990) proposes a series of tests, for both performance and robustness, that may be used to compare controller strategies. The author compares the performance of regulator loops on the basis of the minimisation of the IAE criterion, declaring that such a criterion also minimises peak deviation. The control of time constant dominant processes and time delay dominant processes are considered; generally, the author concludes that model based time delay controllers (such as the Smith predictor) have higher performance than PID controllers for all processes, if the former are specifically tuned to minimise the IAE criterion. However, model based time delay controllers are, in general, less robust than PID controllers; thus, the authors conclude that the preferred controller for a particular task depends on the type of process to be controlled and the relative importance given to performance and robustness. If better performance is to be achieved, the author suggests a feedforward/feedback control strategy. Finally, if the performance of a feedback controller is acceptable, but its robustness is not acceptable, then the author suggests that self-tuning is appropriate.

Overall, the parameter optimised controllers (and specifically the PID controller) are appropriate for the compensation of non-dominant time delay processes. For a process with a dominant time delay, one possibility is to convert the process to a non-dominant time delay process (perhaps by using a Smith predictor), and then design a parameter optimised controller in this Smith predictor structure.

5.3 Structurally optimised controllers

Structurally optimised controllers refer to controllers in which the controller structure and the controller parameters are adapted optimally to the structure and parameters of the process model. This section of the chapter concentrates on the use of such controllers to compensate processes with time delays. The compensators will be discussed under the following headings:

1. The Smith predictor, and its variations.
2. Direct synthesis methods, which are typically based on designing the controller to meet a required output specification; pole placement controllers are an example.
3. Optimal controller design methods, which may be based on a minimum variance or linear quadratic control strategy.
4. Predictive controllers and
5. Other compensation strategies for processes with time delays.

5.3.1 The Smith predictor and its variations

5.3.1.1 Introduction

Smith (1957) defined a method (subsequently entitled a Smith predictor) for the control of a process with a time delay. The method involves effectively removing the time delay from the control loop; an appropriate controller may then be designed for the delay free portion of the process i.e. a controller C^* is designed for a process $G_p e^{-s\tau_p}$ such that the desired closed loop transfer function of the system (in servo mode) is $\frac{G_c G_p}{1+G_c G_p} e^{-s\tau_p}$ (with C^*, G_c and G_p being functions of the Laplace variable, s).

Therefore,

$$\frac{C^* G_p e^{-s\tau_p}}{1+C^* G_p e^{-s\tau_p}} = \frac{G_c G_p}{1+G_c G_p} e^{-s\tau_p} \tag{5.6}$$

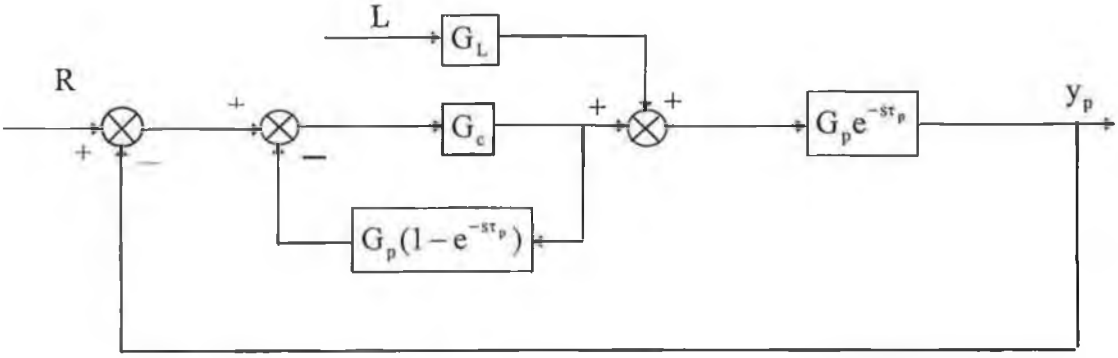
i.e.

$$C^* G_p e^{-s\tau_p}(1+G_c G_p) = G_c G_p e^{-s\tau_p}(1+C^* G_p e^{-s\tau_p}) \tag{5.7}$$

i.e. $$C^{*} = \frac{G_c G_p e^{-s\tau_p}}{G_p(1+G_cG_p)e^{-s\tau_p} - G_cG_p^{\,2}e^{-2s\tau_p}} = \frac{G_c}{1+G_cG_p(1-e^{-s\tau_p})} \qquad (5.8)$$

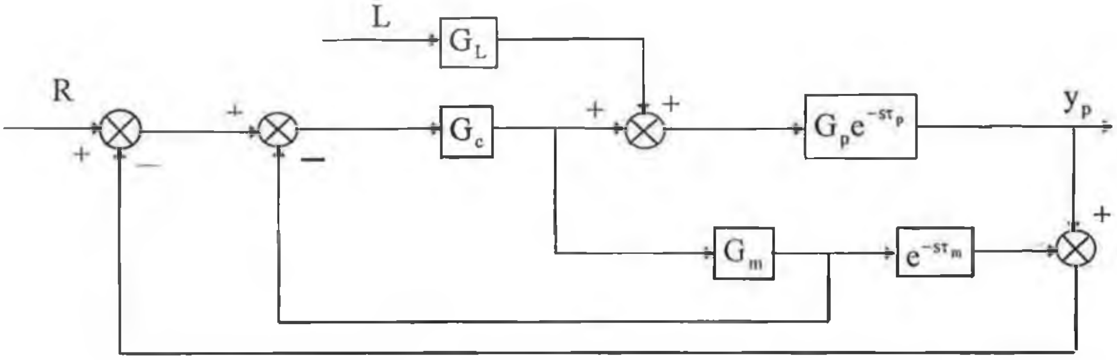
Thus, the implementation of this controller in block diagram form is shown in Figure 5.2 (G_L is a function of the Laplace variable, s).

Figure 5.2: Smith predictor implementation



An alternative implementation is shown in Figure 5.3 (this implementation also recognises that G_p and τ_p are modelled by G_m and τ_m).

Figure 5.3: Alternative Smith predictor implementation



The closed loop servo transfer function of this system is

$$\frac{y_p}{R} = \frac{\dfrac{G_c G_p e^{-s\tau_p}}{1 + G_c G_m (1 - e^{-s\tau_m})}}{1 + \dfrac{G_c G_p e^{-s\tau_p}}{1 + G_c G_m (1 - e^{-s\tau_m})}} \tag{5.9}$$

i.e.

$$\frac{y_p}{R} = \frac{G_c G_p e^{-s\tau_p}}{1 + G_c G_m + G_c \left[G_p e^{-s\tau_p} - G_m e^{-s\tau_m} \right]} \tag{5.10}$$

Therefore, for ideal time delay compensation, the mismatch term $G_c(G_p e^{-s\tau_p} - G_m e^{-s\tau_m})$ must equal zero.

Meyer *et al.* (1976), Astrom and Zhou (1981), Horowitz (1983) and Seborg *et al.* (1989) discuss the applicability of the Smith predictor, especially compared to the PI or PID controller. The latter authors, for instance, quote studies that declare that the performance of the Smith predictor for set point changes can be as much as 30% better than a conventional PID controller based on minimising an ISE criterion (which is the least conservative of the integral criteria). It is also suggested that the Smith predictor can provide an improvement over PI control if the model parameters are within about 30% of the process parameters. Palmor and Blau (1994) suggest that a properly tuned Smith predictor performs much better than a PID controller in loops typical of the process industries, even though the model used in the Smith predictor may be of much lower order than the process; the authors suggest that a FOLPD process model is adequate, with the primary controller in the Smith predictor being of PI or PID structure.

The Smith predictor may be derived from other time delay compensator strategies. Palmor (1982), for instance, shows that the constrained minimum output variance controller for a process with a time delay, if the process and disturbance models are of low order, is a Smith predictor with a PI or PID primary controller. Other authors also show that the Smith predictor may be interpreted as an optimal controller for time delay compensation under certain conditions (Kleinman (1969), Donoghue (1977), Cook and Price (1978), Grimble (1979), Hammerstrom and Waller (1980), Watanabe and Ito (1981), Clark (1985), Durbin (1985)). In summary, it is shown that the optimal controller for a process with a time delay is a Smith predictor for servo applications, or when the disturbance may be considered to be a step input, if the optimal controller is designed using a constrained minimum output variance control law. If the disturbance is not a step input, then the optimal controller may be specified for regulator applications by the inclusion of an appropriate dynamic element in the feedback path of the Smith predictor structure. The design of such dynamic elements is considered in more detail in Chapter 6.

In other contributions, Astrom and Wittenmark (1984) and Landau (1995) show that a Smith predictor (in the z domain) may be derived from the pole placement compensator design for a time delayed process; Middleton and Goodwin (1990) also show this relationship in the delta domain. It is interesting that Soeterboek (1992) states that the Smith predictor (implemented in the discrete time domain) is the best d+1 step ahead predictive controller (d = process time delay index) for a process with time delay, with a constant or random walk measurement disturbance (the author claims that such disturbances are frequently found in industrial applications). In an interesting comment, Morari and Zafiriou (1989) remark that the Smith predictor is an extreme form of lead compensator.

The Smith predictor has been investigated in many simulation and implementation studies. Singh and McEwan (1976), for instance, consider the implementation of the Smith predictor compensated system in a laboratory case study, in which the delay in the predictor is approximated by a second order Pade approximation, realised in continuous time by an appropriate operational amplifier based circuit. The authors show that the servo response of the closed loop compensated system is significantly better than if a PI controller is used, despite the presence of some mismatch between the process and the model. Other such studies have been done by Parrish and Brosilow (1985), Schneider (1988), Papageorgiou and Messner (1989) and Foss and Wasbo (1994). Other contributions that are of interest are those of Shinskey (1990), Hagglund (1992) and Rad *et al.* (1995) (who discuss in detail a time delay compensator that is a special case of the Smith predictor, called the predictive PI (PIP) controller), Young *et al.* (1990) (who control a stable non-minimum phase process using a 'pseudo-predictor', which the authors describe as a Pade approximation for a Smith predictor) and Tan and De Keyser (1994) (who consider the use of neural network based Smith predictors to compensate a non-linear process with a large time delay).

5.3.1.2 The design of the Smith predictor in continuous time

In real applications, it is inevitable that the model will not be a perfect representation of the process, perhaps because the process and the model are of different structure or because the process parameters change in an unknown way with operating conditions. This difference between the process and the model is referred to as 'mismatch'. The presence of mismatch means that perfect time delay compensation using the Smith predictor is not possible; compensation is a particular problem if the process parameters vary. Two approaches are possible to improve the operation of the Smith predictor in these circumstances: the model parameters could be adaptively updated as the process parameters vary, or a robust Smith predictor could be designed, if a limit on the variation of the process parameters is known.

Adaptive model parameter estimation schemes have been implemented (by Marshall (1979), (1980), Bahill (1983), Malik-Zafarei and Jamshidi (1987) and O' Connor (1989), amongst others) to allow the model parameters to track the process parameters. Kaya and Scheib (1984) also propose the tracking of slowly varying parameters of a FOLPD process model (using the method defined by Marshall (1979)); a Smith predictor, with a primary controller in PID form whose parameters are calculated by minimising the ITAE criterion, is then adjusted corresponding to the changes in the model parameter values. Other authors that implement adaptive estimation schemes include Liu (1990), Hang *et al.* (1994b), (1995) and Palmor and Blau (1994); such schemes are also discussed in Chapter 7. The difficulty with many adaptive approaches is that the closed loop system may be unstable as a result of the mismatch, before the model parameters are updated to the process parameters. Therefore, a fundamental requirement is that the Smith predictor compensated system should stay stable in the presence of mismatch (this may be considered a robust stability criterion). The performance of the compensated system in the presence of mismatch (i.e. the fulfilment of a robust performance criterion) is also of interest.

The conditions for stability in the presence of mismatch may be calculated analytically in both the time and frequency domains, by tracking the poles of the closed loop transfer function for changes in the process parameters in the time domain, or by showing that the magnitude of the open loop transfer function is less than 1 when the phase lag equals 180^0, in the frequency domain, as the process parameters change.

Unfortunately, numerical methods are required in both cases to calculate the controller parameters needed to keep the controlled system stable. In addition, knowledge of the process parameters is required; it may be possible to design a (cautious) primary controller if the maximum mismatch between the process and the model parameters is known. Some authors consider creating a deliberate mismatch between the process and model time delays to improve stability (Vit (1979), Marshall and Salehi (1982), Hocken *et al.* (1983)), though it does appear that the process time delay must be known *a priori*.

An alternative is to specify robust stability and robust performance requirements for the Smith predictor implementation, in the presence of mismatch between the process and model parameters. Palmor (1980), (1982), Garcia and Morari (1985), Laughlin and Morari (1987), Laughlin *et al.* (1987), Yamanaka and Shimemura (1987), Gorecki *et al.* (1989), Morari and Zafiriou (1989), Fisher (1991), Santacesaria and Scattolini (1993), Shu *et al.* (1994) and Lee *et al.* (1996) discuss these issues in detail. Laughlin and Morari (1987) and Laughlin *et al.* (1987), for instance, define a single multiplicative perturbation to represent the uncertainty in several real parameters; the authors subsequently derive analytical conditions for robust stability and robust performance of the Smith predictor. The authors use the IMC procedure to formulate an appropriate primary controller. An alternative robust stability condition using μ analysis is discussed by Wang and Skogestad (1993). The authors state that it is normally necessary to approximate the time delay by a rational transfer function to synthesise the μ optimal controller; however, if the controller is in a Smith predictor structure, then it is not necessary to approximate the time delay. Therefore, the robust stability conditions to be evaluated are the same as those of the delay-free system, taking into account any delay uncertainty. In an interesting variation of the above strategy, Wang *et al.* (1994) convert the time delay design problem to a delay free one, by modelling the nominal time delay as uncertainty. A simple nominal model without the time delay is the result, allowing the use of the standard robust stability and performance criteria. The controller is designed in terms of the H_∞ norm, using μ synthesis.

5.3.1.3 Smith predictor modifications in the continuous time domain

The Smith predictor strategy is designed with servo applications in mind. Palmor and Blau (1994) suggest that because of this, and since the Smith predictor contains a model in parallel with the process (which means that the open loop poles are excited by disturbances which in some cases may govern the response), the regulator action of the Smith predictor is less effective. If a disturbance L acts on the system (Figure 5.3), then

$$\frac{y_p}{L} = \frac{G_L\left[1 + G_c G_m\left(1 - e^{-s\tau_m}\right)\right]}{1 + G_c G_m\left(1 - e^{-s\tau_m}\right) + G_c G_p e^{-s\tau_p}} \tag{5.11}$$

A number of modifications on the basic Smith predictor structure have been defined, to improve the disturbance regulator properties of the compensated system. Watanabe and Ito (1981) and Watanabe *et al.* (1983), for example, modify the basic Smith predictor by including a lead-lag compensator in the feedback path of the major loop to make the controlled system less sensitive to disturbances; the method is developed to compensate a SISO process and is extended to compensate MIMO processes. Other such modifications are proposed by Marshall (1979), Kantor and Andres (1980), Palmor and Powers (1985), Wong and Seborg (1986), Romagnoli *et al.* (1988), Wang and Wan (1988), Gorecki *et al.* (1989), Huang *et al.* (1990), Mitchell (1990), Astrom *et al.* (1994), Benouarets and Atherton (1994), Palmor and Blau (1994) and Dastych (1995). These modifications are outlined in Chapter 6 and are discussed in detail by O'Dwyer (1996f).

5.3.1.4 The control of unstable processes using time delay compensators

The control of unstable processes with time delays, using an appropriate compensator, has been considered by Furukawa and Shimemura (1983), amongst others; these authors show that such a process may not be stabilised with a Smith predictor, as the poles of the compensated closed loop system always contain those of the unstable process. De Paor (1985) designs a modified Smith predictor and associated primary controller for the control of an unstable process modelled by an n^{th}

order model plus time delay with one unstable pole. The author factorises the denominator polynomial in the model into a Hurwitz polynomial and a term that corresponds to the part of the process denominator that includes the unstable pole. The author then designs the controller, so that the overall compensated system is asymptotically stable for a range of values of the delay, while satisfying a disturbance rejection criterion. The method amounts to reflecting, in the imaginary axis, the pole in the right half plane. Other time delay compensation strategies for unstable SISO processes with time delay are defined by De Paor (1989), while De Paor and Egan (1990) develop a sampled data control scheme based on De Paor (1989). Ichikawa (1985) and Wang *et al.* (1988) discuss finite spectrum assignment algorithms, and Zheng *et al.* (1995) describes a variable structure controller, to compensate unstable processes with time delays.

Compensation methods have also been defined to control unstable MIMO processes with time delays. Uraz and Ozturk (1985), for instance, propose a predictor control scheme to stabilise unstable multivariable processes containing time delays in either the control or the output variables. The control scheme is composed of a predictor and a compensator that is formed from feedforward and feedback elements. Other such strategies are defined by Jerome and Ray (1992) and Pandiscio and Pearson (1993).

5.3.1.5 The implementation of the Smith predictor in discrete time

The discussion has concentrated thus far on the implementation of Smith predictors in the continuous time domain. However, it is difficult to generate the equivalent of the model time delay using analogue hardware; it is more straightforward to implement a time delay in the discrete time domain (at least if the time delay is an integer multiple of the sample period). Less work appears to have been done into the investigation of the robustness of discrete time Smith predictor implementations, to mismatch between the process and model parameters; an exception is the work done by Palmor and Halevi (1990) and Whalley and Zeng (1994), who propose analytical procedures to investigate the robustness of the Smith predictor, operating under process-model mismatch conditions in the discrete time domain. The procedures proposed are analogous to the procedures proposed by Palmor (1980), (1982) in the continuous time domain.

It is common to estimate the parameters of the process before designing the appropriate primary compensator in the Smith predictor structure. The time delay may be estimated explicitly; such estimation methods are investigated in Chapter 2. An alternative approach is to overparameterise the model; the design of the Smith predictor is based on the model parameters identified, without an explicit estimation of the time delay. One such approach is defined by Chien *et al.* (1985b), who discuss the control of a process with time delay varying between a minimum and a maximum value; the equivalent time delay is put equal to the minimum value, and the extra time delay is subsumed into an overparameterised process model numerator. The model parameters may then be estimated recursively, and a Smith predictor may be used to implement the self-tuning controller strategy. Other such approaches are described by Batur (1986), Wang (1990), Fujikawa and Yamada (1991), Guez and Piovoso (1991) (who design the resulting Smith predictor using neural networks), Mills *et al.* (1991) and Behbehani *et al.* (1993).

Other authors that discuss the implementation of Smith predictors in the discrete time domain include Teng (1990) and Chen and Jong (1993); the latter authors propose the use of an enhanced Smith predictor structure that is composed of a fuzzy model and a discrete time fuzzy filter. Mechanisms for fuzzy model updating and fuzzy controller tuning are applied to reduce the model mismatch and to improve compensated system performance. The procedure appears to improve the robustness of the Smith predictor. Other modifications of the Smith predictor structure in the discrete time domain are discussed by Chotai and Young (1985), (1987), Zhu and Saucier (1992), Li *et al.* (1994) and Landau (1995).

5.3.1.6 The analytical predictor algorithm

Moore *et al.* (1969), Deshpande and Ash (1983), Wong and Seborg (1986) and Seborg *et al.* (1989) discuss the analytical predictor algorithm, which is a discrete time compensator design approach that combines good regulation behaviour with time delay compensation. Fundamentally, the analytical predictor algorithm includes a disturbance filter in the feedback path; it thus has close similarities with methods that are used to improve the regulator performance of Smith predictors. Wong and Seborg (1986) show that the analytical predictor algorithm and the Smith predictor algorithm are equivalent, if the process and model are identical and the compensated system is optimised for servo applications.

Wong and Seborg (1986), Wellons and Edgar (1987) and Seborg *et al.* (1989) outline a modified version of the algorithm, called the generalised analytical predictor algorithm. If the load is assumed to be constant over N sample periods, it may be shown that exact time delay compensation is achieved when the disturbance filter $= z^N$. Such a filter must be approximated, as it involves the implementation of the controller variable N time steps ahead; one such approximation is provided by Seborg *et al.* (1989). The authors declare that, because of the similarity between the IMC structure and the generalised analytical predictor structure, the IMC methodology may be used to calculate an appropriate primary controller design.

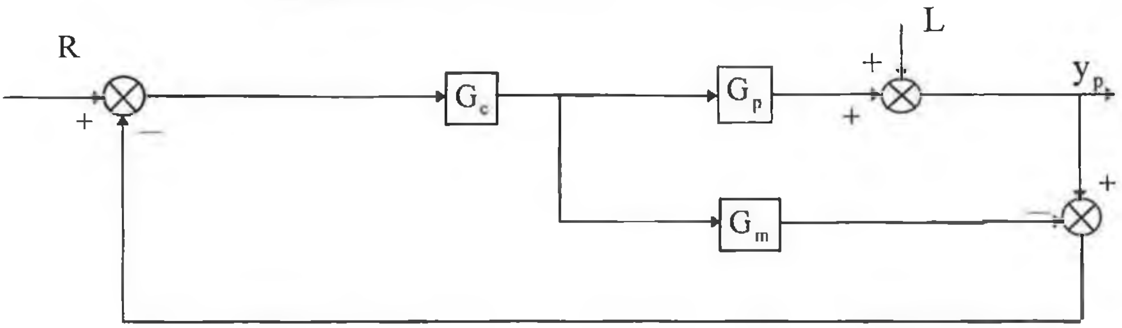
The analytical predictor and generalised analytical predictor algorithms may also be used to compensate MIMO processes with time delays; Huang and You (1994) discuss such an application.

5.3.1.7 The use of the Internal Model Control (IMC) strategy

Seborg *et al.* (1989) and Morari and Zafiriou (1989) discuss the design of the IMC strategy in detail. The block diagram for the IMC structure is set up as shown in Figure 5.4. It may be shown that

$$y_p = \frac{G_c G_p}{1+G_c(G_p - G_m)} R + \frac{1 - G_c G_m}{1+G_c(G_p - G_m)} L \tag{5.12}$$

Figure 5.4: Block diagram for the IMC structure



The design strategy is the same as that given for the design of a robust PID controller in Section 5.2.1.2.5.

The IMC structure is clearly analogous with that of the Smith predictor. Vandeursen and Peperstraete (1995) state that the IMC structure is a generalisation of the modified Smith predictor control structure of Watanabe *et al.* (1983). Thompson (1993) also suggests that the Smith predictor controller is a subset of the internal model controller.

Morari and Zafiriou (1989) outline a procedure for the robust design of a time delay compensator based on the IMC structure. The authors implement the method for the control of a FOLPD process, with a bounded variation in the process time delay. Morari (1987), Thompson (1993) and Datta and Ochoa (1996) also discuss robustness issues in detail.

The IMC methodology may also be used for the design of discrete time delay compensators; this is discussed in detail by Zafiriou and Morari (1985), Seborg *et al.* (1989), Shahrokhi and Naimpour (1992), Peebles *et al.* (1994) and Vandeursen and Peperstraete (1995). In an interesting application, Hunt and Sbarbaro (1991) and Hunt *et al.* (1992) suggest the use of neural networks, in an IMC structure, to implement the adaptive control of non-linear processes.

The compensation of MIMO process models with time delays, using the IMC approach, is described by Garcia and Morari (1985), Luo *et al.* (1992), Wu and Tseng (1992) and Wu *et al.* (1994).

5.3.1.8 Generalised Smith predictors for MIMO process models

Generalised Smith predictors have been defined by a number of authors to control MIMO process models with time delays. Ogunnaike and Ray (1979) and Ray (1981) propose both a discrete time and a continuous time multivariable, multidelay compensator for this application. This compensator is a multivariable version of the Smith predictor; the authors show that the controller reduces to the Smith predictor (and the analytical predictor) for a SISO application. Jerome and Ray (1986) describe the Generalised MultiDelay Compensator (GMDC), which is an expansion of the Smith predictor to control MIMO processes (and a generalisation of the algorithm of Ogunnaike and Ray (1979)); both time delay compensation and MIMO interaction compensation are achieved in a single design. The design of generalised Smith predictors for these applications is also discussed by Alevisakis and Seborg (1974), Donoghue (1977), Hammerstrom and Waller (1980), Watanabe and Sato (1984), Chien *et al.* (1985a), Ozturk and Fardanesh (1991), Triantafyllou and Grosenbaugh (1991), Austin *et al.* (1993) and Desbiens *et al.* (1996). The robustness of the Smith predictor control scheme, when applied to the control of MIMO processes, is discussed by Owens and Raya (1982a), (1982b), Palmor and Halevi (1983), Chu and Wu (1986) and Feng (1991); the latter author derives a sufficient condition for the practical stability of linear multivariable processes, controlled using the Smith predictor strategy.

5.3.2 Direct synthesis methods

5.3.2.1 Introduction

A direct synthesis algorithm is one that results in a controller that gives a specified closed loop response. Direct synthesis methods may be used to specify the controller parameters of low order controllers, such as PI, PID or lead-lag controllers (see Section 5.2.1.2.4); the technique covers pole placement controllers and controllers that achieve a specified closed loop transfer function. For processes that include a time delay, low order controllers are typically specified by approximating the time delay term by an appropriate rational polynomial. This model will thus approximate the process at lower frequencies. This section of the chapter will discuss the design of controllers in which the time delay is typically not approximated. The controller design techniques may be based on a state space or input-output process model.

5.3.2.2 Continuous time domain

Seborg *et al.* (1989) outline the direct synthesis method of controller design in the continuous time domain, for processes that include time delays; the desired closed loop transfer function must include a time delay greater than the process time delay. It may be shown that the controller designed using the method has a time delay compensator structure. If the process is of FOLPD structure, for example, then

$$G_p(s) = \frac{K_p e^{-s\tau_p}}{1 + sT_p} \qquad (5.13)$$

If the desired closed loop response is

$$\left(\frac{C}{R}\right)_d = \frac{K_d e^{-s\tau_d}}{1 + sT_d}, \quad \tau_d \geq \tau_p \qquad (5.14)$$

Then

$$G_c(s) = \frac{1}{G_p(s)} \frac{(C/R)_d}{1 - (C/R)_d} = \frac{1 + sT_p}{K_p e^{-s\tau_p}} \frac{K_d e^{-s\tau_d}}{1 + sT_d - K_d e^{-s\tau_d}} \qquad (5.15)$$

No direct equivalence may be drawn between this structure and the Smith predictor structure. As with the Smith predictor, some difficulties would arise in the practical implementation of this structure (in the continuous time domain). The IMC design strategy (Section 5.3.1.7) has some similarities to the direct synthesis method described; Seborg *et al.* (1989) show that the appropriate choice of the desired closed loop response (direct synthesis method) and the IMC filter (IMC method) can mean that the compensators designed using both approaches are identical.

Direct synthesis methods are also discussed by Sanathanan and Quinn (1987) and Lilja (1990); the former authors calculate an appropriate low order controller for a high order process with time delay, by matching the frequency response of the compensated system with the desired frequency response, based on an appropriate reference model.

5.3.2.3 Discrete time domain

Direct synthesis methods for calculating an appropriate sampled data controller, based on an input-output model approach, have been discussed by many authors; Dahlin (1968), for instance, derives an appropriate controller by assuming that the desired closed loop transfer function is the discrete equivalent of a continuous FOLPD model. The time constant for the closed loop system may be adjusted to give more sluggish control if the process parameters are not known accurately. Such algorithms have been discussed in detail by Chiu *et al.* (1973a), Palmor (1982), Zafiriou and Morari (1985), Seborg *et al.* (1986), Leffew *et al.* (1987), Seborg *et al.* (1989), Dumont (1990), Elnagger *et al.* (1992), (1993) and Dumont *et al.* (1993). Seborg *et al.* (1989) show that the controller implementation may have the disadvantage of allowing intersample ripple in the controlled variable and the controller output.

A pole placement design approach is outlined by Isermann (1989), (1991) and Wellstead and Zarrop (1991), for application to the control of a general process with time delay; Isermann (1991), however, suggests that a large computational effort is associated with the design procedure. Pole placement controllers may be designed by including the time delay in an overparameterised process model; Wellstead *et al.* (1979), Vogel and Edgar (1982), Prasad *et al.* (1985), Seborg *et al.* (1986), and Wang (1990) discuss this approach. However, the order of the polynomial increases as the time delay increases, making the method unattractive for the design of compensators for processes with large time delays (Seborg *et al.* (1986)).

Alternatively, the time delay may be approximated by a rational polynomial; Stahl and Hippe (1987), De Souza *et al.* (1987), (1988) and Salgado *et al.* (1988) design pole placement controllers for the process on this basis. Interestingly, De Souza *et al.* (1987) implement the pole placement controller in the δ (or Euler) discrete domain. However, the robustness of the compensated system at higher frequencies would need to be considered, as the time delay may be poorly modelled at such frequencies (Stahl and Hippe (1987)); interestingly, Hang and Chin (1991) state that extensive simulation experience suggests that as long as most of the time delay is explicitly modelled, then the residual time delay may be safely estimated as a non-minimum phase zero.

Other authors that use pole placement compensators in their applications include Astrom and Zhou (1981), Lammers and Verbruggen (1985), Astrom and Hagglund (1988), Kristinsson and Dumont (1992), Gendron *et al.* (1993), Astrom *et al.* (1993), Chen *et al.* (1994), Kotob *et al.* (1994) and Lundh and Astrom (1994). Other direct synthesis methods are defined by Isermann (1989) and Isermann *et al.* (1992), who discuss the design of "cancellation" controllers (in which part of the controller cancels the poles and zeroes of the process) and "deadbeat" controllers (in which a finite settling time is required for both the controlled variable and the manipulated variable).

Adaptive pole placement controllers (or model reference adaptive controllers (MRAC)) may also be designed to compensate processes with time delays. The design of a MRAC in the continuous time domain for process models with time delay is discussed by Nagarajan and Sajed (1994). More attention appears to have been paid to the design of such controllers in the discrete time domain; for instance, Isermann *et al.* (1992) discuss the design of such a MRAC. The authors state that the discrete time compensated systems designed are stable if the process time delay index is known, the reference model time delay index is greater than or equal to the process time delay index, an upper bound on the process order is known and the process is of minimum phase form. Gawthrop (1977), Barthal and Shin (1993), Meyn and Brown (1993), and Kimura *et al.* (1994) also discuss the design of such compensators in the discrete time domain. The design of a MRAC for MIMO process applications is discussed by Mizuno and Fujii (1983).

State-space design approaches are an alternative to the input-output model approach for direct synthesis compensator design. Such approaches are described by Isermann (1989), Bartolini and Ferrara (1992) and Tsai *et al.* (1994). Isermann (1989) designs the controller for a SISO process model assuming no time delay, and suggests that the time delay compensator design may be facilitated by including the time delay in the system matrix. It is suggested that controllers designed using a pole placement state space approach are recommended for processes of high order, processes with a large time delay or if the process model is known precisely.

Manitius and Olbrot (1979) discuss the finite spectrum assignment approach for the control of linear processes with time delays in the state and/or output variables. The procedure involves designing a feedback law based on pole assignment. This is a time domain approach; a corresponding frequency domain approach that facilitates the arbitrary placing of the finite poles of the process is outlined by Ichikawa (1985). Wang *et al.* (1988) state that the finite spectrum assignment method removes all of the delay from the closed loop characteristic equation (as does the Smith predictor); the method allows the resulting closed loop system to have a finite number of poles located at an arbitrarily preassigned set of points in the complex plane. This means that the method may be used to compensate unstable processes with time delays (unlike the Smith predictor). Manitius (1984) describes a number of control laws, based on an application of a finite spectrum assignment approach, for the control of a process with a time delay. The author shows good closed loop transient responses corresponding to the use of one of the control laws developed; he also states that on-line updating of the controller parameters is possible, as the calculation of the coefficients in the control law is straightforward. In a very recent paper, Yao *et al.* (1996) design observers for processes with time delays using a finite spectrum assignment method. The finite spectrum assignment method is also discussed by Furukawa and Shimemura (1983), Ichikawa (1985), Watanabe and Ouchi (1985), Ortega and Lozano (1988), Wang and Chen (1988) and Wang *et al.* (1988), (1995b).

5.3.2.4 Direct synthesis controller design methods for MIMO process models

Continuous time and discrete time approaches to the design of direct synthesis compensators, using an input-output model approach or a state-space approach, have been defined for the compensation of MIMO processes with time delays. Continuous time compensators using an input-output model approach have been defined by Perng and Ju (1991) and Agamennoni *et al.* (1992); the latter authors design the compensator so that the frequency response of the controller satisfies the criteria that the slow disturbance modes should be rejected by each loop, and that an appropriate dominant time constant for each loop should be attained. A continuous time compensator for MIMO process models is designed using a state-space approach by Kocijan and Korba (1991). Discrete time compensators using an input-output model approach have been explored in more detail; Wellstead *et al.* (1979), Prager and Wellstead (1980), Chai (1986), Lang *et al.* (1986), Seborg *et al.* (1986), Kinneart *et al.* (1987), Kinneart and Hanus (1988), Vogel and Edgar (1988), Isermann (1991), Yamamoto *et al.* (1991), Isermann *et al.* (1992), Mo and Bayoumi (1993), Song and Hardt (1994) and Teng *et al.* (1994) discuss these techniques. Seborg *et al.* (1986) state that the advantages of the MIMO pole placement controller are similar to those of the SISO pole placement controller (i.e. it may be applied to control non-minimum phase processes, it may be detuned to avoid excessive control action, it provides time delay compensation and it facilitates the control of processes with variable time delay); in addition, it allows for the control of MIMO processes with different time delays between the input-output combinations. However, since the method involves overparameterisation of the numerator polynomial, it is not particularly attractive for the control of processes with large time delays. Finally, the design of state-space direct synthesis controllers in the discrete time domain has received some attention; Isermann (1991) discusses the design of such compensators.

5.3.3 Optimal controller design methods

5.3.3.1 Introduction

This section of the chapter will consider the development of optimal design techniques to specify compensator parameters for the control of processes with time delays. The process is assumed to be a time varying linear process, and process and measurement noise may be present. The controller is synthesised to minimise an appropriate criterion; Astrom and Wittenmark (1984) discuss a number of optimal control criteria. One such criterion minimises the variance of the controlled variable (the minimum variance (MV) control strategy); in the discrete time domain, the criterion is

$$J = E[y^2(k+1+d)] \tag{5.16}$$

with d = time delay index. Another criterion minimises the expected value of the square of the controlled variable plus a multiple times the square of the control signal (the linear quadratic (LQ) control strategy or the linear quadratic Gaussian (LQG) control strategy, if Gaussian stochastic disturbances are allowed in the system); in the discrete time domain, the criterion is

$$J = E[y^2(k+1+d) + \rho u^2(k)] \tag{5.17}$$

5.3.3.2 Input-output design approach

General LQ design strategies in the continuous time domain for processes with time delays are discussed by Grimble (1979) and Semino and Scali (1994); the latter authors, for example, calculate the parameters in the weight function that guarantee the required robustness of the response using the IMC design strategy.

Optimal controllers may also be designed in the discrete time domain; Astrom and Wittenmark (1984), (1989), Landau (1990), Isermann (1991) and Isermann *et al.* (1992), for instance, discuss in detail the design of a MV and LQ optimal controller for a process with time delay. Wellstead and Zarrop (1991) declare it is necessary that an accurate value of the process time delay index should be used in the compensation. The authors state that if the model time delay index is less than the process time delay index, then instability may result; if the model time delay index is greater than the process time delay index, then the regulation error is unnecessarily large. Isermann (1991) and Isermann *et al.* (1992) extend the above techniques to the design of LQG and MV compensators for MIMO processes with time delays.

The design and implementation of MV compensators is discussed in detail by Wong and Bayoumi (1982), Clough and Park (1985), Liu and Gertler (1987), Koivo *et al.* (1988), Hu *et al.* (1988), Xu (1988), Ren (1993) and Al-Chalabi and Khalil (1994). The MV control strategy may also be used to design compensators for MIMO processes with known time delays; Borison (1979) and Chien *et al.* (1985) discuss these applications of the method in detail. It is possible to associate the MV controller strategy with some of the compensation strategies discussed earlier; Palmor (1982) and Dumont *et al.* (1993), for example, state that the direct synthesis controller derived by Dahlin (1968) is a minimum variance compensator. Seborg *et al.* (1986) state that the MV controller may be interpreted as a PID controller in addition to a component containing memory of previous control actions, and additional terms providing a form of time delay compensation. In interesting comments on the applicability of the MV algorithm, the same authors declare that a MV controller is not suitable for the control of processes with time delay, because the time delay is likely to be a non-integer multiple of the sample period, which may result in poor, and even unstable, performance. The reason for this is explained by Wellstead and Zarrop (1991), who point out that the zero, due to the modelling in the z domain of the time delay that is not an integer multiple of the sample period, moves along the negative real axis of the z plane, as the "fractional" value of the time delay varies from zero to the sample time. For small values of the "fractional" time delay, the process model will be in non-minimum phase form. A process controlled by a MV controller may be destabilised in these circumstances, as the implementation of the compensator involves the inversion of the model numerator polynomial. Therefore, if stable compensation is to be achieved, the time delay used in the MV controller design must be larger than the actual value of the process time delay (which results in a non-optimal controller). In a similar manner, time varying time delays also cause problems for the design of a MV controller.

The design and implementation of LQ compensators is discussed in detail by Kurz (1979), Kurz and Goedecke (1981), Durbin (1985), Chien *et al.* (1985b), Allidina *et al.* (1985), Seborg *et al.* (1986), Voss *et al.* (1988), Chen and Zhang (1990), Zhang and Chen (1990), Roy *et al.* (1991b), (1991c), (1993a), Pratt and Downing (1994), Yusof *et al.* (1994) and Weerasooriya and Phan (1995). Clark (1985) shows that the implementation of the LQ controller for the control of processes with time delays may be considered to be similar to a Smith predictor strategy, in that the predicted rather than the current measured output is fed back. The author argues that the method takes account of disturbances and may control open loop unstable processes (unlike the Smith predictor). The LQ control strategy may also be used to design compensators for MIMO processes with delays; El-Bagoury and Bayoumi (1980), Koivo (1980), Hahn *et al.* (1982), Dugard *et al.* (1984), Scattolini (1986), Chai (1988), Tade *et al.* (1988a), (1988b), Chai and Ma (1990), Chai (1990), Chen *et al.* (1991), Chai and Wang (1992) and Yin and Asbjornsen (1993) discuss this application of the method in detail.

An interesting feature of the LQ or MV method is that since the process input influences the process output at a number of steps equal to the time delay index later, the minimisation of the output variance, for instance, requires a controller exhibiting 'k step ahead prediction', where k is set equal to the time delay index. Thus, these optimal strategies are intrinsically predictive; predictive controller strategies are discussed in Section 5.3.4.

5.3.3.3 State-space design approach

Continuous time and discrete time approaches have been defined for the specification of optimal controllers for processes with time delays. Anderson and Moore (1989), for example, demonstrate that an optimally designed LQ state controller for a process with time delay is stable provided $\tau < \pi / 3\omega_r$, ω_r = frequency at which the magnitude of the process is unity and τ = time delay. Grimble (1980), Malik-Zafarei and Jamshidi (1987), Palanisamy *et al.* (1988), Gorecki *et al.* (1989), Dadebo and Luus (1992), Dadebo and McAuley (1993), Chyung (1993) and Paraskevopoulos and Samiotakis (1994) discuss in detail the optimal control of SISO processes with time delay, in the continuous time domain, by minimising an LQ cost function, while Malik-Zafarei and Jamshidi (1987) discuss the optimal control of MIMO processes with time delay, using the approach.

Linear quadratic optimal controllers may also be defined in the discrete time domain, using the state-space approach. Joshi and Kaufman (1975), Knobbe (1989), Middleton and Goodwin (1990) and Ha *et al.* (1993) consider the design of the optimal controller for a process model with a time delay; interestingly, Middleton and Goodwin (1990) design the compensator in the δ domain. The LQ strategy may also be used to design compensators for MIMO process models; Ray (1981) presents a general formulation of the controller problem for the control of MIMO processes with time delays. The author discusses in detail the optimal control of such processes in the state space domain, using the LQ control strategy; the material presented is based on the work of Ray and Soliman (1970), who discuss general fundamental results on the optimal control of time delay systems.

5.3.3.4 Other optimisation strategies for SISO process models

Other optimisation strategies may also be used for compensator design; one such approach is the time optimal controller design method that involves the determination of the control function which drives the system to a desired state in minimum time. Algorithms based on determining this controller for a process with time delay in the continuous time domain are defined by Latour *et al.* (1967), Malik-Zafarei and Jamshidi (1987) and Lin *et al.* (1993). Other optimisation approaches are defined in the discrete time domain by Ozbay and Peery (1993) and Lublinsky and Fradkov (1993).

5.3.4 Predictive controllers

Soeterboek (1992) states that predictive controllers calculate a future controller output sequence so that the predicted output of the process is "close" to the desired process output. If only the first element of the controller output sequence (i.e. at sample time k) is used to control the process, and the predicted sequence is repeated at sample time k+1, then this is referred to as a "receding horizon" predictive controller. The predictive controller is designed by minimising an appropriate cost function. One such predictive controller is the unified predictive controller, as outlined by Soeterboek (1992). A cost function to be optimised is defined, with appropriate constraints. The author subsequently derives the unified predictive control law; it is shown that, under certain circumstances, the unified predictive control law may be interpreted as a pole placement controller (with prediction). The author also discusses the stability robustness and the performance robustness of the unified predictive control law; the gain margin and time delay margin (which is the change in the model time delay required to drive the controlled system unstable) are used to evaluate the stability robustness of the method. The author states that the unified predictive controller action may be summarised as the driving of future controlled variables $y(k+i)$ to follow as closely as possible the desired set point input $r(k+i)$ over the next i samples. This is done at each sampling instant k by calculating the future set point sequence $r(k+i)$, predicting the controlled variable at sample k+i, minimising an appropriate cost function to provide the suggested control sequence $u(k+i)$, implementing the first element of the control sequence $u(k+1)$ and then repeating the calculations.

Soeterboek (1992) shows that the Dynamic Matrix Control (DMC) algorithm, the Predictive Control Algorithm (PCA), the Model Algorithmic Control (MAC) algorithm, the Generalised Predictive Control (GPC) algorithm (of Clark *et al.* (1987a), (1987b)), the Extended Prediction Self-Adaptive Control (EPSAC) algorithm and the Extended Horizon Adaptive Control (EHAC) algorithm (of Ydstie (1984)) may be regarded as special cases of the unified predictive control algorithm discussed above. The author states that all of these controllers can compensate processes with time delays successfully.

In a review paper, Kwon (1994) compares the model predictive control (MPC) strategy, the GPC strategy and the receding horizon control (RHC) strategy as applied to the control of linear unconstrained processes, linear constrained processes and non-linear processes. These strategies are compared in historical origin, in models and cost indices used, in predictive strategies used and in the solutions obtained to the predictive control problem. In addition, the stability and robustness properties of the controlled systems are explored, when the relevant algorithms are implemented. A large number of industrial applications of the strategies are reported.

The GPC strategy is a subset of the unified predictor controller strategy. The GPC strategy, and relevant implementations, is discussed in detail by Clark *et al.* (1987a), (1987b), Dumont (1990), Demircioglu and Gawthrop (1991), Fisher (1991), Isermann *et al.* (1992), Willis *et al.* (1992) (who uses neural networks to implement the strategy), Camacho and Bourdons (1993), Dumont *et al.* (1993), Elnagger *et al.* (1993), Jolly *et al.* (1993) and Yamamoto *et al.* (1994). Camacho and Bourdons (1993), for instance, model the process as a FOLPD model, and apply the GPC strategy to calculate a control sequence that minimises a multistage cost function. The authors provide simple formulae to calculate the appropriate tuning parameters required.

Other predictive controller strategies are discussed by Ydstie (1984) (who considers the use of the EHAC algorithm), Zevros and Dumont (1988), Garcia *et al.* (1989), Ishida and Zhan (1992), (1993), Chisci and Mosca (1993) (who considers the use of the RHC algorithm), Rad and Lo (1994), Rangaiah and Krishnaswamy (1994), Zhu and Ling (1994), Maiti *et al.* (1994) (who considers the use of a DMC algorithm), Austin *et al.* (1994) and Huang and You (1995).

5.3.5 Other compensation strategies for processes with time delays

5.3.5.1 Feedforward control

Feedforward control (or more realistically, feedforward-feedback control) may be used to compensate a process with time delay. If a disturbance acts on the controlled variable, a feedforward element with transfer function equal to the reciprocal of the process may be used to eliminate the effect of the disturbance. However, as Isermann (1991) suggests, ideal feedforward control is not possible if the process has a time delay or if the zeroes of the (discrete) process model or disturbance model are on or outside the unit circle. A feedforward controller is typically used in association with a feedback controller structure.

Isermann (1991) outlines a number of methods for the design of a feedforward-feedback controller for a process with time delay in the discrete time domain, based on the use of a cancellation controller strategy, the use of a parameter optimised feedforward controller strategy, the use of a state variable technique and the use of a LQG controller design method. A disadvantage of feedforward control is that accurate knowledge of the process is required. An adaptive feedforward-feedback controller implementation, in which the model parameters are continuously updated, may be used to facilitate wider use of the strategy; one such algorithm is discussed by Isermann (1991).

Other authors that discuss the design and implementation of feedforward-feedback controllers for the control of processes with time delay are Palmor and Powers (1985) (who use a feedforward controller in conjunction with a Smith predictor), Peter and Isermann (1988), Rao and Perdikaris (1988), Hagglund and Astrom (1989), (1991), Morari and Zafiriou (1989), Newell and Lee (1989), Shinskey (1990) (who suggests that feedforward-feedback control offers the capability of decreasing the IAE of the compensated closed loop system by another order of magnitude over an appropriate PID feedback controller), Astrom (1991), Mills *et al.* (1991), Astrom *et al.* (1992) and Chyung (1993); Astrom *et al.* (1992), for instance, suggest that if the process may be modelled by a FOLPD model, and if the conditions (i) $\tau_m / T_m \le 0.6$ and (ii) the product of the ultimate gain and the process gain is greater than 2.25 are not fulfilled, then feedforward-feedback compensation may be appropriate.

A feedforward element, in addition to the feedback controller, may also be used to compensate MIMO processes with time delays; Uraz and Ozturk (1985), Chai (1990), Chai and Ma (1990) and Gawthrop and Nomikos (1990) consider appropriate compensator designs in detail.

5.3.5.2 Other strategies

Many time delay compensation strategies do not naturally fall into the previous categories discussed. Chou *et al.* (1989), for example, design a feedback controller to robustly stabilise an uncertain, saturating process with a time delay. A large number of other strategies, defined in the literature, have been outlined by O'Dwyer (1996c), for the stabilisation of SISO and MIMO process models with time delay; many of these strategies are concerned with the conditions for stability for the closed loop control of processes with time delays.

Finally, the robustness of compensator designs is a topic that has been the focus of increasing attention in recent years. Earlier sections of the chapter have discussed a number of approaches that facilitate the achievement of a certain degree of robustness, when a process with time delay is being controlled. Other authors define other robustness procedures than have been encountered thus far; Lamaire *et al.* (1991), for example, propose a method for robust process estimation in the frequency domain in which bands on the frequency domain estimation error are developed. This procedure allows the on-line design of a robust control law. Other robustness strategies that have appeared in the recent literature are outlined by O'Dwyer (1996c).

5.3.6 Conclusions

This chapter has considered a wide variety of methods for the compensation of processes with time delay, in both the continuous time and discrete time domains. The wide spectrum of methods covered, and the dependence of the choice of compensator method on the application, means that an overall conclusion as to the best method to use is not appropriate. However, it has been concluded in Section 5.2.3 that parameter optimised controllers are not appropriate for the compensation of dominant time delay processes. Structurally optimised compensators are more appropriate for the compensation of such processes, as the controller structure and controller parameters may be adapted optimally to the structure and to the parameters of the process model. It has been decided to concentrate on the development of a modified version of the Smith predictor, with the aim of improving the regulator response of the Smith predictor, while retaining a similar servo response; this decision has been made as many methods, viewing the compensation problem from a variety of perspectives, appear to present the Smith predictor as the optimal (or a component of the optimal) controller for dominant delay processes. An alternative perspective is that the Smith predictor structure may be used to reduce the dominance of the time delay term, and thus facilitate the conversion of the compensation problem from the control of a dominant time delay process to the control of a non-dominant time delay process (though process-model mismatch difficulties do have to be taken into account); the primary controller in the predictor may therefore be designed using a parameter optimised approach.

The development of the modified Smith predictor is detailed in Chapter 6.

CHAPTER 6

The compensation of processes with time delay by using an appropriately modified Smith predictor

6.1 Introduction

It has been concluded in Chapter 5 that the Smith predictor is a component of the optimal controller for dominant time delay processes. However, as discussed in Section 5.3.1.3, the Smith predictor strategy is designed with servo applications in mind. It has therefore been decided to develop a modified version of the Smith predictor, with the aim of improving the regulator response of such a compensator; the development is performed in the continuous time domain.

Initially, the Smith predictor, and modifications of the Smith predictor that appear in the literature, are considered in more detail. A generalised modified Smith predictor is then defined, and the design of such a structure to achieve ideal servo and regulator responses, with the elimination of process-model mismatch, is described. Subsequently, the design of the modified Smith predictor to achieve realistic servo and regulator action is discussed. This latter development involves the approximation of a time advance term; an appropriate approximation is detailed. The servo and regulator responses of the modified Smith predictor and the Smith predictor are then compared for a number of process-model structures. The performances of the Smith predictor and the modified Smith predictor are compared analytically by calculating the sensitivities of the output of the compensated systems to changes in the plant parameters. Finally, appropriate conclusions as to the efficacy of the modified Smith predictor implementation are drawn.

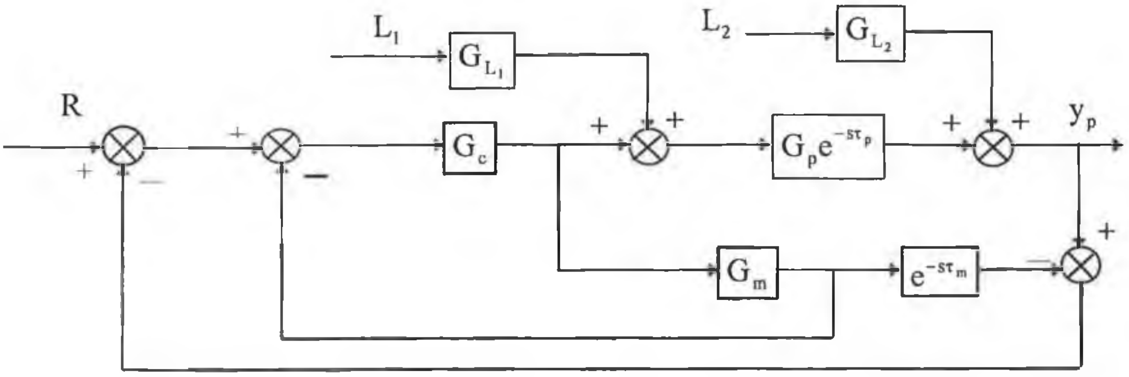
The algorithms and simulations in this chapter are considered in more detail by O'Dwyer (1996f), and are also discussed by O'Dwyer and Ringwood (1996).

6.2 The Smith predictor and its modifications

6.2.1 Introduction

The design of controllers for processes with long time delays has been of interest to academics and practitioners for several decades. In a seminal contribution, Smith (1957) proposed a technique that facilitates the removal of the time delay term from the closed loop characteristic equation. This method, labelled the Smith predictor, has been the subject of numerous experimental and theoretical studies, and is considered in detail in Chapter 5. A block diagram of the Smith predictor structure (with process and measurement noise) is provided in Figure 6.1 (with each of the terms $G_{L_1}, G_{L_2}, G_c, G_p$ and G_m being functions of the Laplace variable, s).

Figure 6.1: Block diagram of the Smith predictor structure



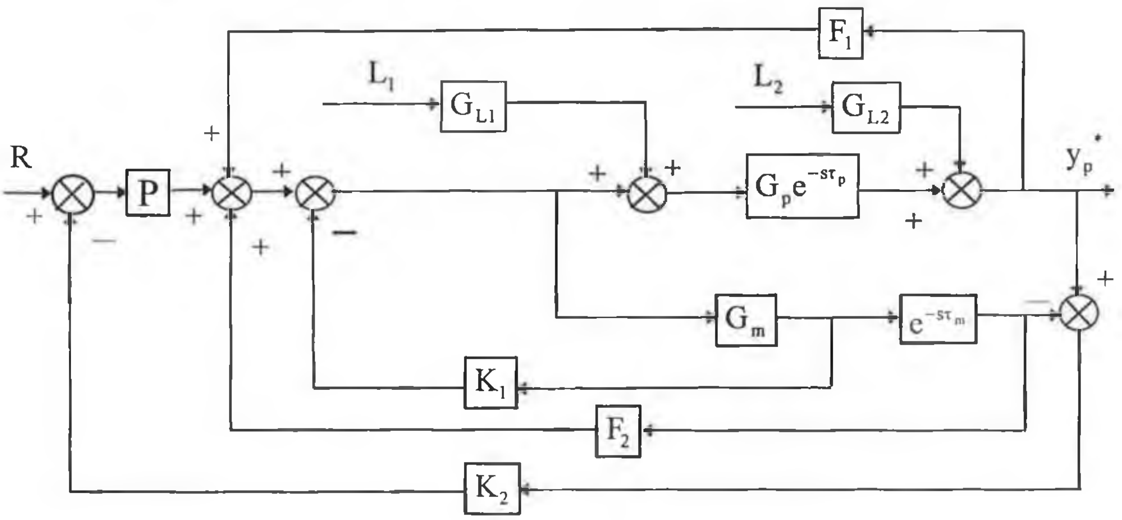
The response of the compensated system is as follows:

$$y_p = \frac{(G_p G_c e^{-s\tau_p})R + (G_{L_1} G_p e^{-s\tau_p} L_1 + G_{L_2} L_2)(1 + G_m G_c[1 - e^{-s\tau_m}])}{1 + G_c G_m + G_c(G_p e^{-s\tau_p} - G_m e^{-s\tau_m})} \qquad (6.1)$$

A number of authors have proposed modifications to the Smith predictor structure to improve the regulator response of the compensated system and/or to reduce the effect on either the servo or the regulator response of process-model mismatch. Some authors (e.g. Marshall (1979), Hammerstrom and Waller (1980), Kantor and Andres (1980), Watanabe and Ito (1981), Marshall and Salehi (1982), Watanabe *et al.*

(1983), Hocken *et al.* (1983), Wong and Seborg (1986), Romagnoli *et al.* (1988), Mitchell (1990), Astrom *et al.* (1994) and Dastych (1995)) suggest that improved responses may be obtained if appropriate dynamic terms are included in either the outer feedback loop or the inner feedback loop of the Smith predictor. Other authors suggest either the inclusion of extra dynamic elements from either the process or model outputs (Benouarets and Atherton (1994)), the inclusion of a dynamic element in the forward path of the compensator between the outer and inner feedback loops (Huang *et al.* (1990)) or the feedforward of a measurable disturbance signal acting on the process through an appropriate dynamic element (Palmor and Powers (1985)). Many of the modifications of the Smith predictor structure discussed are subsets of the implementation provided in Figure 6.2 (with each of the terms P, F_1, F_2, K_1 and K_2 being functions of the Laplace variable, s and y_p^* is the process output of the modified Smith predictor).

Figure 6.2: Block diagram of a generalised Smith predictor structure



The response of the system in Figure 6.2 may be derived to be

$$y_p^* = \frac{(G_p P e^{-s\tau_p})R + \left(1 + G_m\left[K_1 + \{F_2 - K_2 P\}e^{-s\tau_m}\right]\right)\left(G_{L2}L_2 + G_p G_{L1} e^{-s\tau_p} L_1\right)}{1 + \left(G_m K_1 + G_m[F_2 - K_2 P]e^{-s\tau_m} - G_p[F_1 - K_2 P]e^{-s\tau_p}\right)} \qquad (6.2)$$

6.2.2 Optimising the servo and regulator responses

One may optimise the servo and regulator responses, and minimise the effect of the mismatch between the process and the model, by appropriate design of three of the five dynamic elements in Figure 6.2 (as only three specifications need to be fulfilled). This redundancy means that there are ten possible modifications to the Smith predictor structure, outlined in Table 6.1 (with blank terms indicating the dynamic elements to be designed).

Table 6.1: Possible modifications of the Smith predictor structure

Modification	P	F_1	F_2	K_1	K_2
Case 1	1		0		
Case 2	1	0			
Case 3		0			1
Case 4		0		1	
Case 5	1				1
Case 6			0	1	
Case 7		0	0		
Case 8			0		1
Case 9	1			1	
Case 10				1	1

If it is desired to achieve ideal servo and regulator action (i.e. $y_p^*/R = 1.0$, $y_p^*/L_1 = 0.0$ and $y_p^*/L_2 = 0.0$) with mismatch elimination, then the application of equation (6.2) will reveal the requirements on the individual dynamic elements in the modifications. Of the ten modifications, it may be shown by applying equation (6.2) that three of the modifications do not facilitate the achievement of the full specification (Cases 7, 8 and 9 in Table 6.1) and two of the modifications fulfil the specification in an identical manner (Cases 4 and 10). Table 6.2 summarises the requirements on the six remaining cases, to achieve the specification, with Cases 1a to 6a corresponding to Cases 1 to 6 in Table 6.1.

An example calculation clarifies how these requirements are achieved. Taking Case 1, equation (6.2) becomes

$$y_p^{*} = \frac{(G_p e^{-s\tau_p})R + \left(1 + G_m\left[K_1 - K_2 e^{-s\tau_m}\right]\right)\left(G_{L2}L_2 + G_p G_{L1} e^{-s\tau_p} L_1\right)}{1 + \left(G_m K_1 - G_m K_2 e^{-s\tau_m} - G_p\left[F_1 - K_2\right]e^{-s\tau_p}\right)} \tag{6.3}$$

From equation (6.3), ideal servo action is achieved when $G_p e^{-s\tau_p}/(1 + G_m K_1) = 1$ i.e. when

$$K_1 = (G_p e^{-s\tau_p} - 1)/G_m \tag{6.4}$$

From equation (6.3), ideal regulator action is achieved when $1 + G_m(K_1 - K_2 e^{-s\tau_m}) = 0$ i.e. when, substituting from equation (6.4),

$$K_2 = G_p e^{-s\tau_p}/G_m e^{-s\tau_m} \tag{6.5}$$

From equation (6.3), mismatch is eliminated when $G_p(K_2 - F_1)e^{-s\tau_p} = G_m K_2 e^{-s\tau_m}$ i.e. when, substituting from equation (6.5),

$$F_1 = [G_p e^{-s\tau_p}/G_m e^{-s\tau_m}] - 1 \tag{6.6}$$

Table 6.2: Possible methods of achieving ideal servo and regulator action, with elimination of mismatch

Implementation	P	F_1	F_2	K_1	K_2
Case 1a	1	a_1	0	a_2	$a_1 + 1$
Case 2a	1	0	$-a_1$	a_2	1
Case 3a	arbitrary	0	$-Pa_1$	a_5	1
Case 4a	a_3	0	$-a_4$	1	1
Case 5a	1	$-1/(1+a_1)$	$-1-a_1$	a_2	1
Case 6a	a_3	a_4	0	1	$a_1 + 1$

with $a_1 = \frac{G_p}{G_m} e^{-s(\tau_p - \tau_m)} - 1$, $a_2 = \frac{G_p e^{-s\tau_p} - 1}{G_m}$, $a_3 = \frac{1 + G_m}{G_p e^{-s\tau_p}}$, $a_5 = \frac{G_p P e^{-s\tau_p} - 1}{G_m}$ and

$$a_4 = (1 + G_m)\left(\frac{1}{G_m e^{-s\tau_m}} - \frac{1}{G_p e^{-s\tau_p}}\right).$$

However, all of the representations in Table 6.2 require inversion of the model or (unknown) process transfer function, and/or the model or (unknown) process time delay, to set up one or more of the required dynamic elements, which means that the design involves the specification of a non-proper transfer function. Such non-proper transfer functions would need to be approximated; supplementary calculations reveal that any deviation of the parameters from those used in the non-proper specification drives the closed loop modified Smith predictor unstable.

6.2.3 The design of a realistic modified Smith predictor

It is more realistic to design the modified Smith predictor such that the servo response of the compensated system is similar, for example, to that of a FOLPD model, with a corresponding regulator response. Such responses may also be achieved by using the IMC strategy described by Morari and Zafiriou (1989). An appropriate choice of three dynamic elements in each of the cases taken will fulfil the requirements for realistic regulator and servo action (with the servo requirement being equivalent to the response to an appropriately ordered model plus time delay), together with the requirement for mismatch elimination. The requirements on the dynamic elements are summarised in Table 6.3; Cases 1b to 6b correspond to Cases 1a to 6a in Table 6.2, under the new design requirements.

An example calculation clarifies how these requirements are achieved. The realistic servo response is $G_p G_c e^{-s\tau_p} / (1 + G_p G_c)$, with the corresponding regulator response (to a process disturbance) being $G_p G_{L1} e^{-s\tau_p} / (1 + G_p G_c)$; G_c is considered to be an appropriate primary controller in the Smith predictor structure. From equation (6.3), realistic servo action is achieved when $G_p e^{-s\tau_p} / (1 + G_m K_1)$ = $G_p G_c e^{-s\tau_p} / (1 + G_p G_c)$ i.e. when

$$K_1 = (1 + G_pG_c - G_c)/G_cG_m \tag{6.7}$$

From equation (6.3), realistic regulator action is achieved when $1/(1 + G_pG_c)$ = $[1 + G_m(K_1 - K_2e^{-s\tau_m})]/(1 + G_mK_1)$ i.e. when, substituting from equation (6.7),

$$K_2 = 1/e^{-s\tau_p} \tag{6.8}$$

The mismatch is eliminated when the conditions defined in equation (6.6) are fulfilled.

Table 6.3: The design of realistic servo and regulator action, with mismatch elimination

Implementation	P	F_1	F_2	K_1	K_2
Case 1b	1	b_6	0	b_1	b_2
Case 2b	1	0	$-b_6$	b_1	$e^{s\tau_p}$
Case 3b	b_3	0	$-b_6b_3e^{s\tau_p}$	1	$e^{s\tau_p}$
Case 4b	1	b_4	$-b_2$	b_1	1
Case 5b	b_3	$b_6b_3e^{s\tau_p}/G_p$	0	1	b_2
Case 6b	b_3	$b_3(1-e^{s\tau_p})$	$b_3b_5e^{s\tau_m}$	1	1

with $b_6 = \dfrac{G_p}{G_m}e^{-s(\tau_p-\tau_m)} - 1$, $b_1 = \dfrac{1 + G_pG_c - G_c}{G_cG_m}$, $b_2 = \dfrac{G_p}{G_me^{-s\tau_m}}$, $b_3 = \dfrac{G_c(1 + G_m)}{1 + G_pG_c}$,

$b_4 = \dfrac{G_p(e^{-s\tau_p} - 1) - G_me^{-s\tau_m}}{G_pe^{-s\tau_p}}$ and $b_5 = \dfrac{G_me^{-s\tau_m} - G_p}{G_m}$.

Unfortunately, non-proper transfer functions are required to implement the specifications; for example, if G_c is considered to be a PI controller and G_p and G_m are considered to be in first order lag (FOL) form, then for Case 1b, for example, K_1 is specified using a dynamic element of denominator order 1 and numerator order 2, and

K_2 and F_1 are specified using terms involving the inversion of the model and the (unknown) process parameters.

If the specification that the mismatch must be eliminated is relaxed (i.e. if the dynamic elements in the modified Smith predictor are defined assuming the model parameters equal the process parameters), then, from Table 6.3, the terms in Cases 1b and 2b are specified using a dynamic element of denominator order 1 and numerator order 2, and the inversion of the model time delay. An inversion of the model time delay is also required in the terms of Cases 3b, 5b and 6b. Table 6.4 summarises the required design of the dynamic elements for realistic servo action and regulator action (to a process disturbance) assuming the model parameters equal the process parameters (when general forms of G_m and G_c are taken); Cases 1c to 6c correspond to Cases 1b to 6b in Table 6.3.

Table 6.4: The design of realistic servo and regulator action, with the process parameters assumed equal to the model parameters.

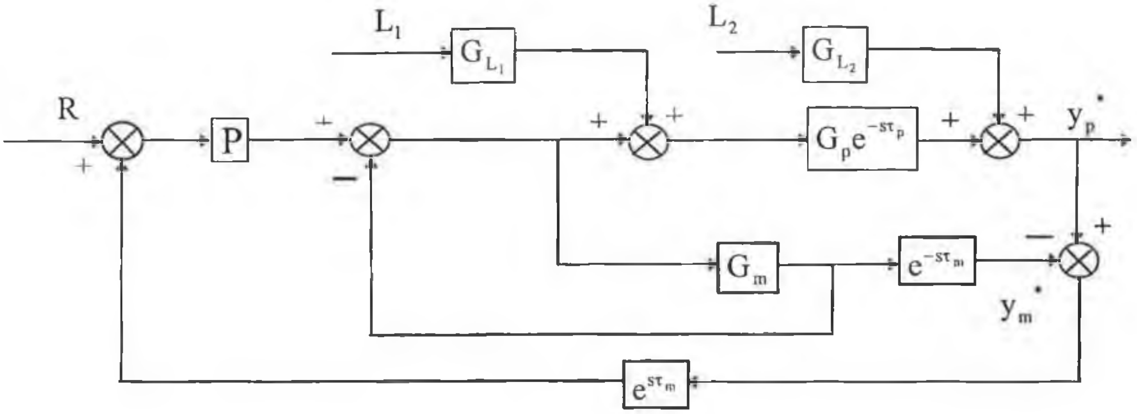
Implementation	P	F_1	F_2	K_1	K_2
Case 1c	1	0	0	c_1	$e^{s\tau_m}$
Case 2c	1	0	0	c_1	$e^{s\tau_m}$
Case 3c	c_2	0	0	1	$e^{s\tau_m}$
Case 4c	1	$-e^{s\tau_m}$	$-e^{s\tau_m}$	c_1	1
Case 5c	c_2	0	0	1	$e^{s\tau_m}$
Case 6c	c_2	$c_2(1-e^{s\tau_m})$	$c_2(1-e^{s\tau_m})$	1	1

with $c_1 = \dfrac{1+G_mG_c-G_c}{G_cG_m}$ and $c_2 = \dfrac{G_c(1+G_m)}{1+G_mG_c}$.

Table 6.4 shows that Cases 3c/5c and 6c are more suitable for design than Cases 1c/2c or 4c as the c_1 term in the latter modifications is non-proper. Of course, in all cases, it is required to specify the (non-proper) inversion of the model time delay.

In summary, the modified Smith predictor design to achieve realistic servo action and regulator action (to a process disturbance) may be represented in block diagram form as shown in Figure 6.3 (corresponding to Case 3c/5c, Table 6.4); a realistic implementation involves appropriately approximating the time advance term. The method has some similarities with that implemented by Huang *et al.* (1990), though the block diagram structures in both designs are different. The method also has similarities with the extended Smith predictor defined by Marshall (1979); in particular, the methods defined by Watanabe and Ito (1981) and Watanabe *et al.* (1983), in which a lead-lag compensator is placed in the feedback path of the outer loop of the Smith predictor to improve its regulator response, are structurally similar to the defined method.

Figure 6.3: Modified Smith predictor design considered



It is now necessary to approximate the time advance term. An interesting approximation is provided by Huang *et al.* (1990), as follows:

$$e^{\tau_m s} \approx \frac{1 + B(s)}{1 + B(s)e^{-s\tau_m}} \tag{6.9}$$

with $B(s) = k/(1 + Ts)$. Other, less accurate, approximations are suggested by Hocken *et al.* (1983) (who use a time delay equal to the difference between the process and model time delays) and Romagnoli *et al.* (1988) (who use a lag controller dynamic element).

The time advance approximation may be improved by defining $B(s)$ as a phase lead network i.e.

$$B(s) = (as + 1)/(as + p), \quad p > 1 \tag{6.10}$$

From Figure 6.3, the responses of the modified Smith predictor, to R and L_1, using the time advance approximation (equation (6.9)), are as follows:

$$\frac{y_p^*}{R} = \frac{G_p G_c e^{-s\tau_p}}{1 + G_c G_m + G_c \left[\frac{1 + B(s)}{1 + B(s) e^{-s\tau_m}} \right] \left(G_p e^{-s\tau_p} - G_m e^{-s\tau_m} \right)} \qquad (6.11)$$

and

$$\frac{y_p^*}{L_1} = \frac{G_p e^{-s\tau_p} G_{L1} \left[1 + G_m G_c \left(1 - e^{-s\tau_m} \right) + B(s) e^{-s\tau_m} \right]}{\left(1 + G_c G_m \right) \left(1 + B(s) e^{-s\tau_m} \right) + G_c \left(1 + B(s) \right) \left(G_p e^{-s\tau_p} - G_m e^{-s\tau_m} \right)} \qquad (6.12)$$

The responses in equations (6.11) and (6.12) reduce to the servo response and the regulator response (to a process disturbance) of the Smith predictor (equation (6.1)) when the approximation for the time advance term is unity.

The time advance approximation in equation (6.9) facilitates the achievement of the zero steady state offset requirement for the servo and regulator responses. From Figure 6.3, the open loop servo response (in an equivalent unity gain negative feedback control system) is

$$G_{OL}^{SERVO} = \frac{G_p G_c e^{-s\tau_p}}{\left(1 + G_c G_m \right) + G_c \left[\frac{1 + B(s)}{1 + B(s) e^{-s\tau_m}} \right] \left(G_p e^{-s\tau_p} - G_m e^{-s\tau_m} \right) - G_p G_c e^{-s\tau_p}} \qquad (6.13)$$

and the open loop regulator response (to a process disturbance) is

$$G_{OL}^{REGULATOR} = \frac{G_p e^{-s\tau_p} \left[1 + G_c G_m \left(1 - e^{-s\tau_m} \right) + B(s) e^{-s\tau_m} \right]}{\left(1 + G_c G_m \right) \left(1 + B(s) e^{-s\tau_m} \right) + G_c \left(1 + B(s) \right) \left(G_p e^{-s\tau_p} - G_m e^{-s\tau_m} \right) - G_p e^{-s\tau_p} X}$$

(6.14)

with $X = 1 + G_c G_m \left(1 - e^{-s\tau_m} \right) + B(s) e^{-s\tau_m}$. The steady state error of the closed loop compensated system to a step input is

$$e_{ss} = \lim_{s \to 0} \frac{sR(s)}{1+G_{OL}} = \lim_{s \to 0} \frac{U}{1+G_{OL}} \tag{6.15}$$

$G_{OL} = G_{OL}^{SERVO}$ or $G_{OL}^{REGULATOR}$, as appropriate, and U = step amplitude. The substitution of equations (6.13) and (6.14) into equation (6.15) reveals that the steady state offset equals zero, for both servo and regulator responses to a step input.

6.2.4 The design of the time advance approximation

A systematic approach for the design of the time advance approximation is difficult to directly deduce from equations (6.11) or (6.12), or indeed from equations (6.13) or (6.14), because of the complex manner in which the phase lead network, B(s), is incorporated into all these equations. It was decided to recommend a design procedure based on simulation work. The seven simulated processes considered in Chapter 4 (equations (4.57) to (4.63)) and their FOLPD and SOSPD models (calculated using the two stage frequency domain method described in Chapter 4) were simulated (in SIMULINK), in the modified Smith predictor structure, for the purpose of specifying a design procedure. The following design procedure for B(s) was determined to be appropriate.

1. The value of a (equation (6.10)) is chosen equal to the time constant of the FOLPD model of the process.
2. The value of p (equation (6.10)) is chosen iteratively to (a) ensure servo and regulator system stability over the full range of variation of the process parameters anticipated and (b) facilitate a better regulator response (and a similar servo response) to that achieved by the original Smith predictor. The starting value of p is chosen as 10.
3. The time delay in the time advance approximation (denominator of equation (6.9)) is chosen as the FOLPD process model time delay.

6.3 Simulation results

The performance of the modified Smith predictor in Figure 6.3, with the time advance approximation designed using the procedure detailed in Section 6.2.4, has been compared with the performance of the corresponding Smith predictor, in the SIMULINK environment. The parameters of the processes considered were allowed to vary about the nominal process parameter values (which were provided by equations (4.57) to (4.63), inclusive); the maximum variation of the process parameters is assumed known *a priori*. The PI and PID primary controllers specified for both predictors are robust to the possible process/model mismatches considered.

A number of simulation results, representing the process/model combinations considered, showing the operation of the modified Smith predictor and the Smith predictor are provided in Figures 6.4 to 6.27. A unit step response is applied to obtain the servo and regulator responses. Taking Simulation 1 as an example, the nominal process (equation (4.57)) is $G_p e^{-s\tau_p} = 2e^{-1.4s}/(1+0.7s)$. The model is put equal to the nominal process, with the process variation allowed being between $1.6e^{-1.2s}/(1+0.5s)$ (labelled $G_{p1}e^{-s\tau_{p1}}$) and $2.4e^{-1.6s}/(1+0.9s)$ (labelled $G_{p2}e^{-s\tau_{p2}}$). The primary controller, G_c, is designed to provide a servo time constant of 1.0 seconds, when the process and model parameters coincide. The design procedure in Section 6.2.4 is used to specify B(s); G_c and the value of p in B(s) are chosen together to ensure stability of the modified Smith predictor over the process parameter variations permitted. The primary controller designed is more conservative than that designed for the robust Smith predictor using the IMC design procedure of Morari and Zafiriou (1989). Simulations 2 to 4 show the performance of the modified Smith predictor, with a mismatch between the nominal process and the model; the model (of the nominal process) is calculated using the two stage frequency domain method described in Chapter 4. G_c is designed to fulfil the performance requirements detailed in the simulations (assuming that the (unknown) nominal process is modelled by the FOLPD model) with robustness guaranteed, as in Simulation 1, by the combined effect of G_c and B(s).

The simulation results presented in Figures 6.4 to 6.27 show that the modified Smith predictor tends to facilitate better regulator responses, with similar servo responses, compared to the Smith predictor, if the desired servo response is relatively slow.

Simulation 1: $G_m e^{-s\tau_m} = 2e^{-1.4s}/(1+0.7s) = G_p e^{-s\tau_p}$. $B(s) = (0.7s+1)/(0.7s+10)$. G_c is specified assuming a servo time constant of 1.0s, when the process and model parameters coincide i.e. $G_c = 0.35(1+1/0.7s)$.

(a) $G_{p1} e^{-s\tau_{p1}} = 1.6e^{-1.2s}/(1+0.5s)$

Figure 6.4: Servo response

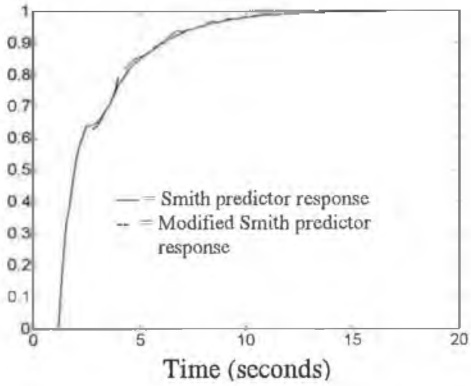
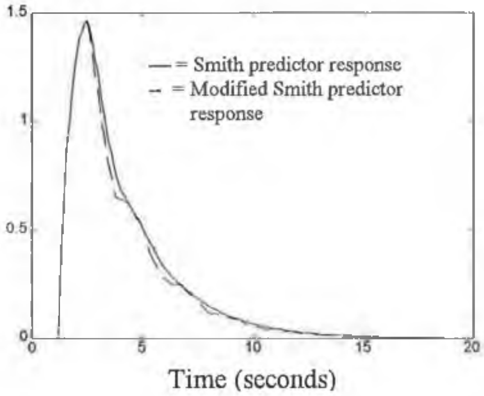


Figure 6.5: Regulator response



(b) $G_p e^{-s\tau_p} = 2e^{-1.4s}/(1+0.7s)$

Figure 6.6: Servo response

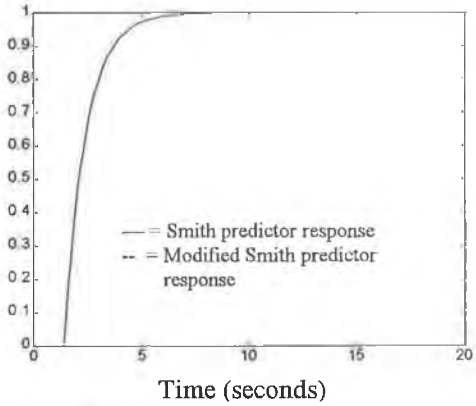
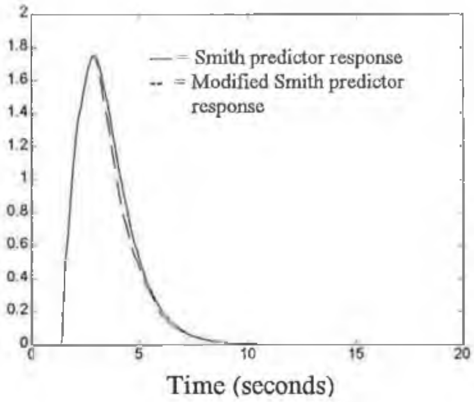


Figure 6.7: Regulator response



(c) $G_{p2} e^{-s\tau_{p2}} = 2.4e^{-1.6s}/(1+0.9s)$

Figure 6.8: Servo response

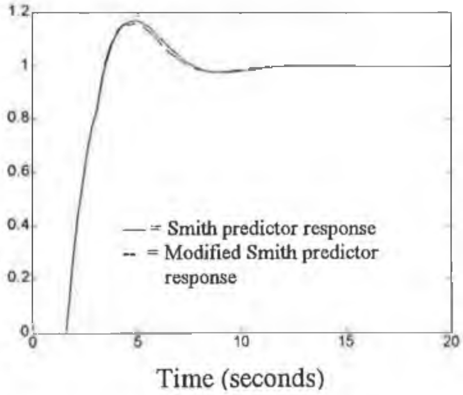
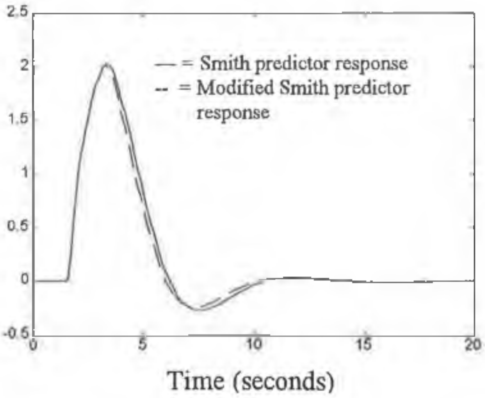


Figure 6.9: Regulator response



Simulation 2: Nominal process $G_p e^{-s\tau_p} = 2e^{-s}/1+8.5s+22.5s^2+18s^3$. FOLPD model $G_m e^{-s\tau_m} = 1.82e^{-3.47s}/(1+7.68s)$ with $B(s) = (7.68s+1)/(7.68s+20)$. G_c is specified assuming a servo time constant of 2.0s, when the process and model are identical i.e. $G_c = 2.11(1+1/7.68s)$.

(a) $G_{p1}e^{-s\tau_{p1}} = 1.2e^{-0.7s}/1+5.9s+15.7s^2+12.6s^3$

Figure 6.10: Servo response

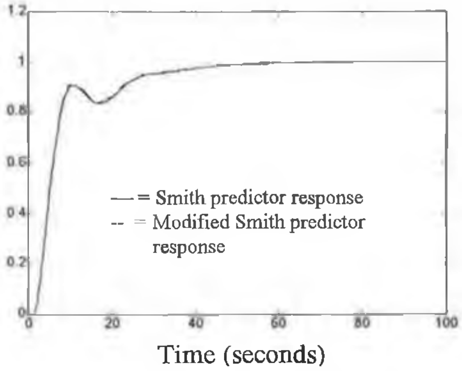
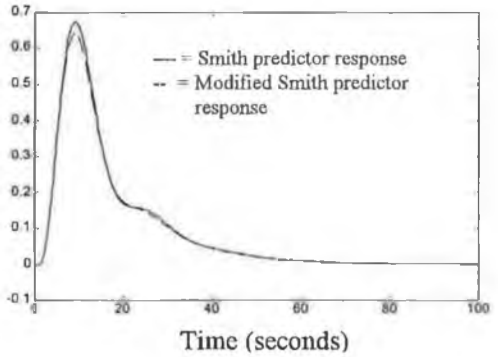


Figure 6.11: Regulator response



(b) $G_p e^{-s\tau_p} = 2e^{-s}/1+8.5s+22.5s^2+18s^3$

Figure 6.12: Servo response

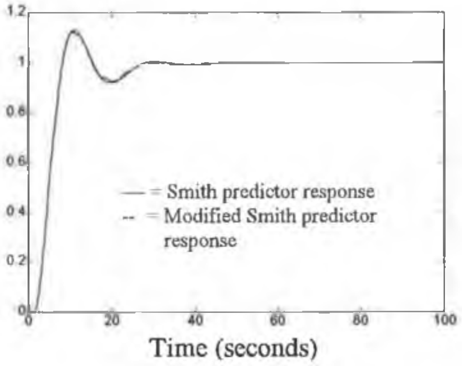
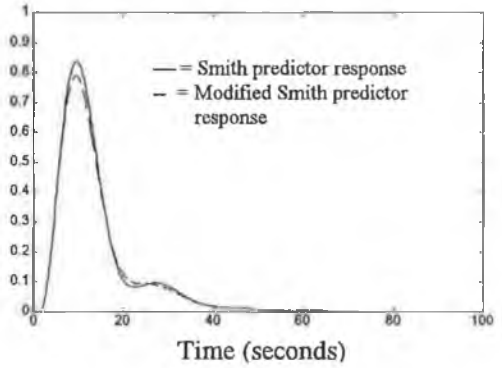


Figure 6.13: Regulator response



(c) $G_{p2}e^{-s\tau_{p2}} = 2.8e^{-1.3s}/1+11s+29.3s^2+23.4s^3$

Figure 6.14: Servo response

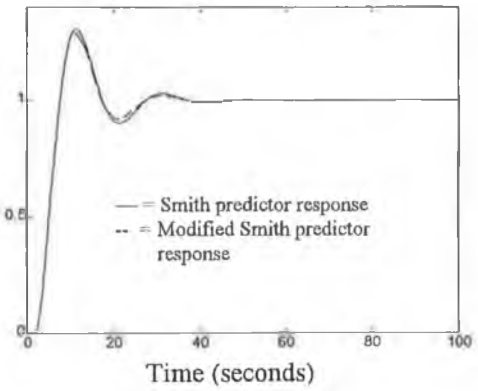
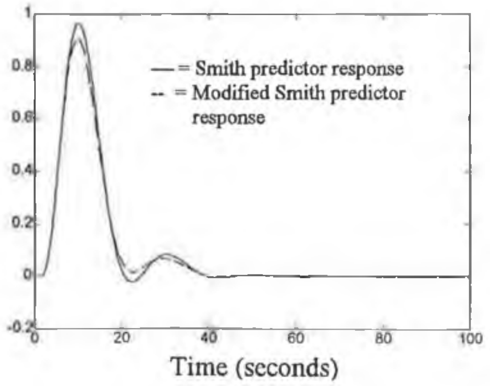


Figure 6.15: Regulator response



Simulation 3: Nominal process $G_p e^{-s\tau_p} = 2e^{-s}/1 + 8.5s + 22.5s^2 + 18s^3$. FOLPD model $G_m e^{-s\tau_m} = 1.82e^{-3.47s}/(1 + 7.68s)$. $B(s) = (7.68s + 1)/(7.68s + 40)$. G_c is specified by setting the damping factor of the closed loop system response equal to 1, when the process and model are identical i.e. $G_c = 2.42(1 + 1/6.97s + 2.03s/(1 + 0.37s))$.

(a) $G_{p1} e^{-s\tau_{p1}} = 1.2e^{-0.7s}/1 + 5.9s + 15.7s^2 + 12.6s^3$

Figure 6.16: Servo response

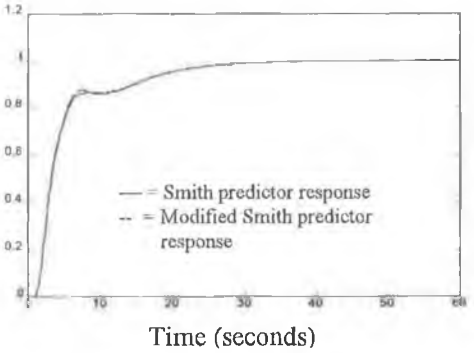
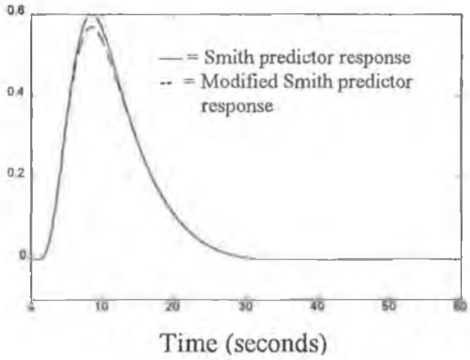


Figure 6.17: Regulator response



(b) $G_p e^{-s\tau_p} = 2e^{-s}/1 + 8.5s + 22.5s^2 + 18s^3$

Figure 6.18: Servo response

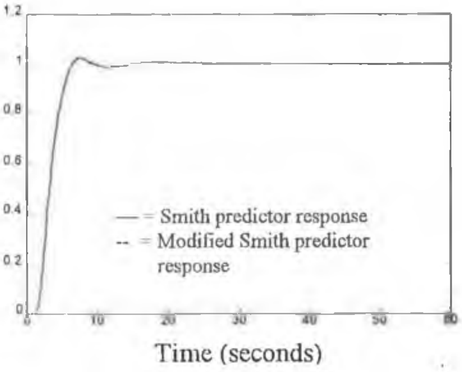
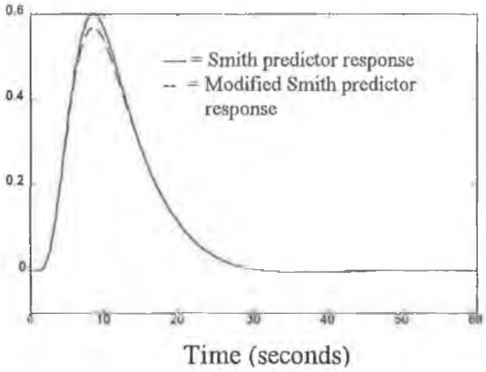


Figure 6.19: Regulator response



(c) $G_{p2} e^{-s\tau_{p2}} = 2.8e^{-1.3s}/1 + 11s + 29.3s^2 + 23.4s^3$

Figure 6.20: Servo response

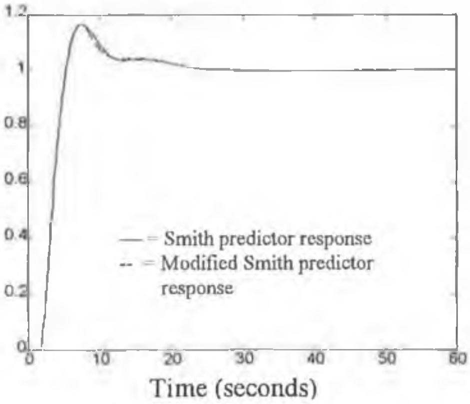
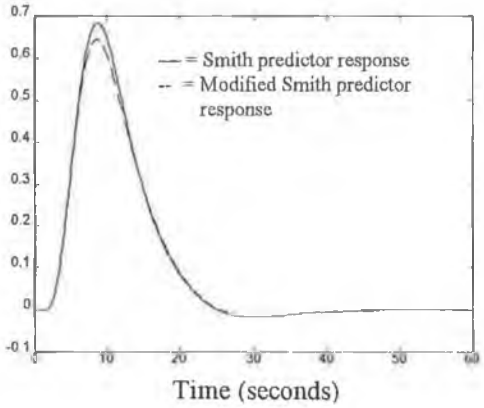


Figure 6.21: Regulator response



Simulation 4: Nominal process $G_p e^{-s\tau_p} = 2e^{-s}/1+8.5s+22.5s^2+18s^3$. FOLPD model $G_m e^{-s\tau_m} = 1.82e^{-3.47s}/(1+7.68s)$. $B(s) = (7.68s+1)/(7.68s+10)$. G_c is specified by optimising the ITAE criterion (in servo mode), when the process and model are identical i.e. $G_c = 0.647(1+1/4.78s)$.

(a) $G_{p1} e^{-s\tau_{p1}} = 1.2e^{-0.7s}/1+5.9s+15.7s^2+12.6s^3$

Figure 6.22: Servo response

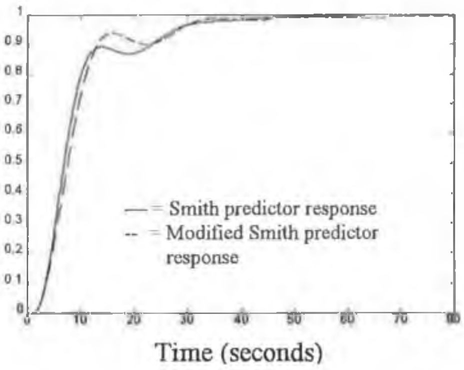
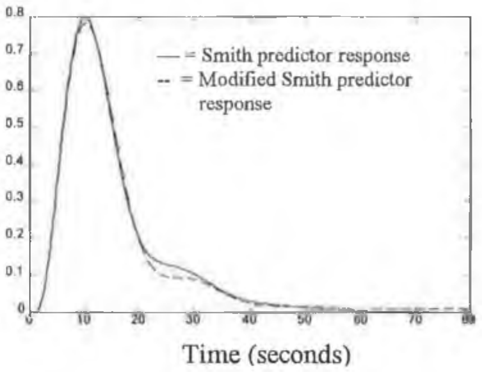


Figure 6.23: Regulator response



(b) $G_p e^{-s\tau_p} = 2e^{-s}/1+8.5s+22.5s^2+18s^3$

Figure 6.24: Servo response

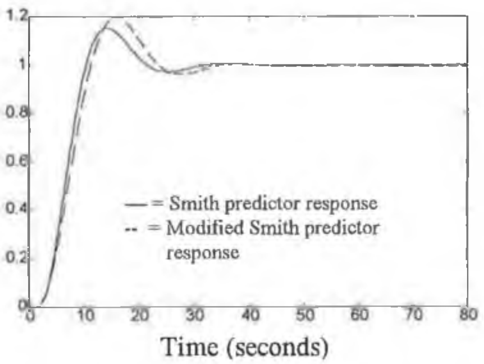
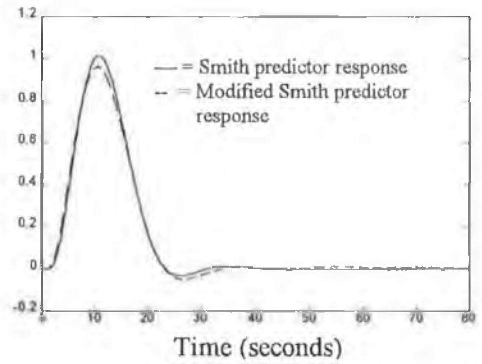


Figure 6.25: Regulator response



(c) $G_{p2} e^{-s\tau_{p2}} = 2.8e^{-1.3s}/1+11s+29.3s^2+23.4s^3$

Figure 6.26: Servo response

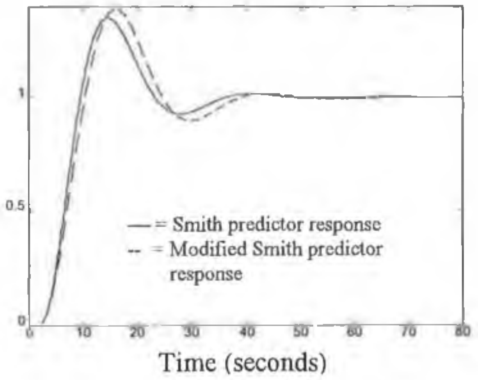
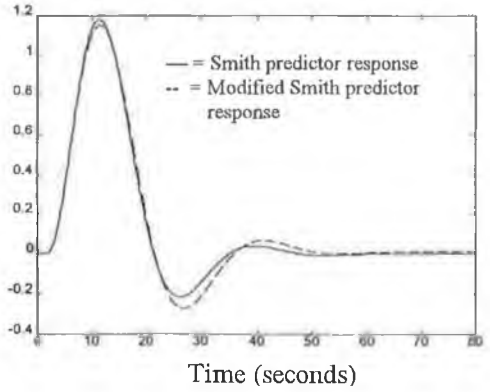


Figure 6.27: Regulator response



6.4 Sensitivity analysis

The simulation results have shown that the modified Smith predictor provided a modest improvement in the servo and regulator responses over the Smith predictor. The sensitivities of the outputs of the Smith predictor and the modified Smith predictor to (small) changes in the process parameters are also of interest. The sensitivity of a transfer function, T, to changes in a plant parameter, α, is defined as

$$S_{\alpha}^{T} = \frac{\alpha}{T}\frac{\partial T}{\partial \alpha} \tag{6.16}$$

The servo transfer function of the modified Smith predictor is given by

$$\frac{y_p^*}{R} = \frac{G_p G_c e^{-s\tau_p}}{1 + G_c G_m + G_c\left[\frac{1 + B(s)}{1 + B(s)e^{-s\tau_m}}\right]\left(G_p e^{-s\tau_p} - G_m e^{-s\tau_m}\right)} \tag{6.17}$$

General expressions for the process and model transfer functions are as follows:

$$G_p(s) = \frac{K_p\left(1 + b_{1p}s + \ldots. + b_{jp}s^j\right)e^{-s\tau_p}}{\left(1 + a_{1p}s + \ldots. + a_{ip}s^i\right)} \tag{6.18}$$

and

$$G_m(s) = \frac{K_m\left(1 + b_{1m}s + \ldots. + b_{km}s^k\right)e^{-s\tau_m}}{\left(1 + a_{1m}s + \ldots. + a_{lm}s^l\right)} \tag{6.19}$$

Straightforward calculation using equation (6.17) reveals that

$$\frac{\partial y_p^*}{\partial \tau_p} = \frac{R\left[1 + G_c G_m - G_c G_m K_2 e^{-s\tau_m}\right]\left[-sG_c G_p e^{-s\tau_p}\right]}{\left[1 + G_c G_m + G_c K_2\left(G_p e^{-s\tau_p} - G_m e^{-s\tau_m}\right)\right]^2} = -sT^{MSP}(s)\left[1 - K_2 T^{MSP}(s)\right]R \tag{6.20}$$

with

$$K_2 = \frac{1 + B(s)}{1 + B(s)e^{-s\tau_m}} \tag{6.21}$$

and $T^{MSP}(s)$ is the servo transfer function of the modified Smith predictor. For the Smith predictor (using the servo response portion of equation (6.1)), it may be calculated that

$$\frac{\partial y_p}{\partial \tau_p} = \frac{R\left[1 + G_c G_m - G_c G_m e^{-s\tau_m}\right]\left[-sG_c G_p e^{-s\tau_p}\right]}{\left[1 + G_c G_m + G_c\left(G_p e^{-s\tau_p} - G_m e^{-s\tau_m}\right)\right]^2} = -sT^{SP}(s)\left[1 - T^{SP}(s)\right]R \qquad (6.22)$$

with $T^{SP}(s)$ being the servo transfer function of the Smith predictor. The sensitivities of the servo response for the modified Smith predictor and the Smith predictor, respectively, with respect to changes in the process time delay (using equations (6.16), (6.20) and (6.22)) are

$$S_{\tau_p}^{T^{MSP}} = \frac{\tau_p}{T^{MSP}(s)R}\frac{\partial y_p^*}{\partial \tau_p} = -s\tau_p\left[1 - K_2 T^{MSP}(s)\right] \qquad (6.23)$$

$$S_{\tau_p}^{T^{SP}} = \frac{\tau_p}{T^{SP}(s)R}\frac{\partial y_p}{\partial \tau_p} = -s\tau_p\left[1 - T^{SP}(s)\right] \qquad (6.24)$$

Similar calculations to those done in equations (6.16) to (6.24) show that

$$\frac{\partial y_p^*}{\partial K_p} = \frac{R\left[1 + G_c G_m - G_c G_m K_2 e^{-s\tau_m}\right]\left[G_c G_p e^{-s\tau_p}\right]}{K_p\left[1 + G_c G_m + G_c K_2\left(G_p e^{-s\tau_p} - G_m e^{-s\tau_m}\right)\right]^2} = \frac{T^{MSP}(s)}{K_p}\left[1 - K_2 T^{MSP}(s)\right]R \qquad (6.25)$$

$$\frac{\partial y_p}{\partial K_p} = \frac{R\left[1 + G_c G_m - G_c G_m e^{-s\tau_m}\right]\left[G_c G_p e^{-s\tau_p}\right]}{K_p\left[1 + G_c G_m + G_c\left(G_p e^{-s\tau_p} - G_m e^{-s\tau_m}\right)\right]^2} = \frac{T^{SP}(s)}{K_p}\left[1 - T^{SP}(s)\right]R \qquad (6.26)$$

$$S_{K_p}^{T^{MSP}} = \frac{K_p}{T^{MSP}(s)R}\frac{\partial y_p^*}{\partial K_p} = \left[1 - K_2 T^{MSP}(s)\right] \qquad (6.27)$$

and

$$S_{K_p}^{T^{SP}} = \frac{K_p}{T^{SP}(s)R}\frac{\partial y_p}{\partial K_p} = \left[1 - T^{SP}(s)\right] \qquad (6.28)$$

In a similar manner, the following equations may be calculated:

$$\frac{\partial y_p^*}{\partial b_{xp}}=\frac{R\left[1+G_cG_m-G_cG_mK_2e^{-s\tau_m}\right]\left[G_cG_pe^{-s\tau_p}\right]}{\left[1+G_cG_m+G_cK_2\left(G_pe^{-s\tau_p}-G_me^{-s\tau_m}\right)\right]^2}\frac{s^x}{1+b_{1p}s+....+b_{jp}s^j}$$

$$=\frac{s^x}{1+b_{1p}s+....+b_{jp}s^j}T^{MSP}(s)\left[1-K_2T^{MSP}(s)\right]R,\quad x=1,.....,j \tag{6.29}$$

$$\frac{\partial y_p}{\partial b_{xp}}=\frac{R\left[1+G_cG_m-G_cG_me^{-s\tau_m}\right]\left[G_cG_pe^{-s\tau_p}\right]}{\left[1+G_cG_m+G_c\left(G_pe^{-s\tau_p}-G_me^{-s\tau_m}\right)\right]^2}\frac{s^x}{1+b_{1p}s+....+b_{jp}s^j}$$

$$=\frac{s^x}{1+b_{1p}s+....+b_{jp}s^j}T^{SP}(s)\left[1-T^{SP}(s)\right]R,\quad x=1,.....,j \tag{6.30}$$

$$S_{b_{xp}}^{T^{MSP}}=\frac{b_{xp}}{T^{MSP}(s)R}\frac{\partial y_p^*}{\partial b_{xp}}=\frac{b_{xp}s^x}{1+b_{1p}s+....+b_{jp}s^j}\left[1-K_2T^{MSP}(s)\right] \tag{6.31}$$

and

$$S_{b_{xp}}^{T^{SP}}=\frac{b_{xp}}{T^{SP}(s)R}\frac{\partial y_p}{\partial b_{xp}}=\frac{b_{xp}s^x}{1+b_{1p}s+....+b_{jp}s^j}\left[1-T^{SP}(s)\right] \tag{6.32}$$

Finally, the following expressions may be obtained:

$$\frac{\partial y_p^*}{\partial a_{xp}}=\frac{R\left[1+G_cG_m-G_cG_mK_2e^{-s\tau_m}\right]\left[G_cG_pe^{-s\tau_p}\right]}{\left[1+G_cG_m+G_cK_2\left(G_pe^{-s\tau_p}-G_me^{-s\tau_m}\right)\right]^2}\frac{-s^x}{1+a_{1p}s+....+a_{ip}s^i}$$

$$=\frac{-s^x}{1+a_{1p}s+....+a_{ip}s^i}T^{MSP}(s)\left[1-K_2T^{MSP}(s)\right]R,\quad x=1,.....,i \tag{6.33}$$

$$\frac{\partial y_p}{\partial a_{xp}}=\frac{R\left[1+G_cG_m-G_cG_me^{-s\tau_m}\right]\left[G_cG_pe^{-s\tau_p}\right]}{\left[1+G_cG_m+G_c\left(G_pe^{-s\tau_p}-G_me^{-s\tau_m}\right)\right]^2}\frac{-s^x}{1+a_{1p}s+....+a_{ip}s^i}$$

$$=\frac{-s^x}{1+a_{1p}s+....+a_{ip}s^i}T^{SP}(s)\left[1-T^{SP}(s)\right]R,\quad x=1,.....,i \tag{6.34}$$

$$S^{T^{MSP}}_{a_{xp}} = \frac{a_{xp}}{T^{MSP}(s)R}\frac{\partial y_p^{*}}{\partial a_{xp}} = \frac{-a_{xp}s^x}{1+a_{1p}s+....+a_{ip}s^i}\left[1-K_2T^{MSP}(s)\right] \tag{6.35}$$

and

$$S^{T^{SP}}_{a_{xp}} = \frac{a_{xp}}{T^{SP}(s)R}\frac{\partial y_p}{\partial a_{xp}} = \frac{-a_{xp}s^x}{1+a_{1p}s+....+a_{ip}s^i}\left[1-T^{SP}(s)\right] \tag{6.36}$$

For all parameters, it may be calculated (using equations (6.23), (6.24), (6.27), (6.28), (6.31), (6.32), (6.35) and (6.36)) that the ratio of the sensitivity of the modified Smith predictor to that of the Smith predictor is the same i.e.

$$\frac{S^{T^{MSP}}_{\alpha}}{S^{T^{SP}}_{\alpha}} = \frac{1-K_2T^{MSP}(s)}{1-T^{SP}(s)} \tag{6.37}$$

where α equals K_p, a_{xp}, b_{xp} or τ_p. The magnitude (and phase) of the ratio in Figure (6.37) will vary with frequency. A practical example is used to demonstrate a typical variation; the parameters taken are broadly similar to those indicated in Simulation 1 viz.

$$G_m e^{-s\tau_m} = \frac{2.0e^{-1.4s}}{1+0.7s} \tag{6.38}$$

with

$$(a)\, G_{p1}e^{-s\tau_{p1}} = \frac{1.6e^{-1.2s}}{1+0.5s},\ (b)\, G_p e^{-s\tau_p} = \frac{2.0e^{-1.4s}}{1+0.7s},\ (c)\, G_{p2}e^{-s\tau_{p2}} = \frac{2.4e^{-1.6s}}{1+0.9s} \tag{6.39}$$

and

$$G_c = 0.7\left(1+\frac{1}{0.7s}\right) \tag{6.40}$$

and

$$K_2 = \frac{1+\dfrac{0.7s+1}{0.7s+10}}{1+\dfrac{0.7s+1}{0.7s+10}e^{-1.4s}} \tag{6.41}$$

Figure 6.28 shows that the modified Smith predictor tends to be less sensitive than the Smith predictor to process parameter variations at lower frequencies, but more sensitive than the Smith predictor to process parameter variations at some higher frequencies. If an ideal time advance term is used for K_2, a similar pattern is seen (Figure 6.29).

Figure 6.28: $S_{\alpha}^{T^{MSP}} / S_{\alpha}^{T^{SP}}$ when the parameters defined in equations (6.38) to (6.41) are used.

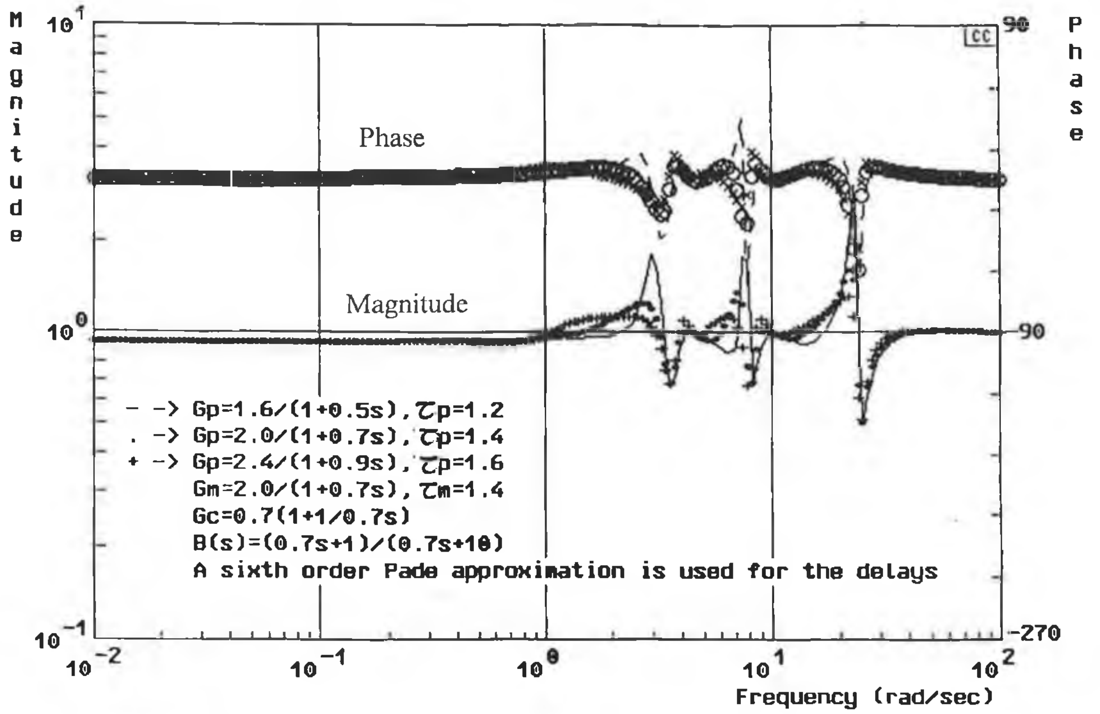
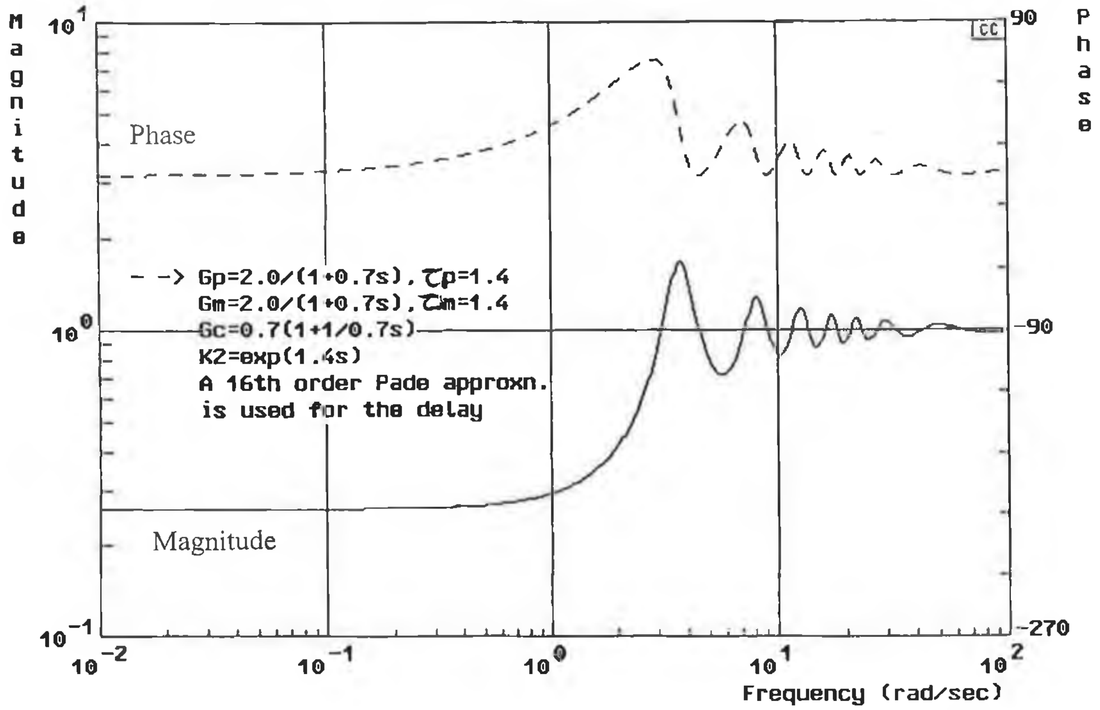


Figure 6.29: $S_{\alpha}^{T^{MSP}} / S_{\alpha}^{T^{SP}}$ when the parameters defined in equations (6.38) to (6.40) are used; the exact value of the time advance is used.



By applying equation (6.37), it may be shown that at low frequencies, for a general process and model structure, and if $G_c = K_c(1+1/T_i s)$, then

$$\frac{S_\alpha^{T^{MSP}}}{S_\alpha^{T^{SP}}} = 1 - \frac{K_c K_m \tau_m}{(1+p)(K_c K_m \tau_m + T_i)} \tag{6.42}$$

with

$$K_2 = \left(1 + \frac{T_m s + 1}{T_m s + p}\right) \Bigg/ \left(1 + \frac{T_m s + 1}{T_m s + p} e^{-s\tau_m}\right) \tag{6.43}$$

with T_m = FOLPD model time constant. This result means that the sensitivity of the modified Smith predictor to process parameter variations is less than that of the Smith predictor at lower frequencies, which conforms with the simulation results seen in Figure 6.28.

It may be shown, by performing similar calculations to those done in equations (6.16) to (6.37), that the ratio of the sensitivities of the modified Smith predictor to the Smith predictor is the same (with respect to the variation of each of the process parameters), regardless of whether the sensitivity of the regulator transfer function or the sensitivity of the servo transfer function is in question. This result applies when the disturbance input is at the input to the process; if the disturbance input is at the output of the process, then similar calculations to those done in equations (6.16) to (6.37) show that

$$\frac{S_\alpha^{T^{MSP}}}{S_\alpha^{T^{SP}}} = \frac{K_2 T^{MSP}(s)}{T^{SP}(s)} \tag{6.44}$$

The sensitivity, to process parameter variations, of the modified Smith predictor in these circumstances equals the sensitivity of the Smith predictor at lower frequencies and tends to be much greater at higher frequencies. Figures 6.30 and 6.31 show this pattern of behaviour, for the simulation conditions taken in equations (6.38) to (6.41), when both an approximation for the time advance term is used (Figure 6.30) and when the exact time advance term is used (Figure 6.31).

Overall, the sensitivity analysis indicates that, at lower frequencies, the modified Smith predictor transfer function is either less sensitive or as sensitive to process parameter variations as is the Smith predictor transfer function.

Figure 6.30: $S_{\alpha}^{T^{MSP}} / S_{\alpha}^{T^{SP}}$ when the parameters defined in equations (6.38) to (6.41) are used, with the disturbance present on the process output.

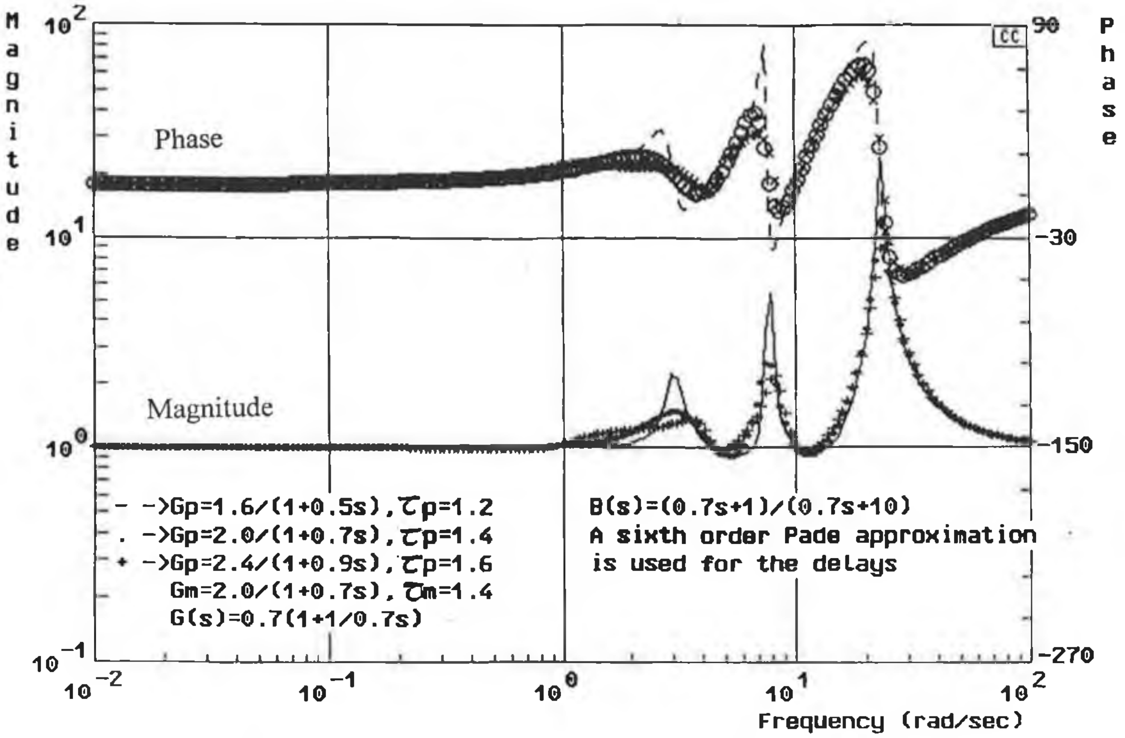
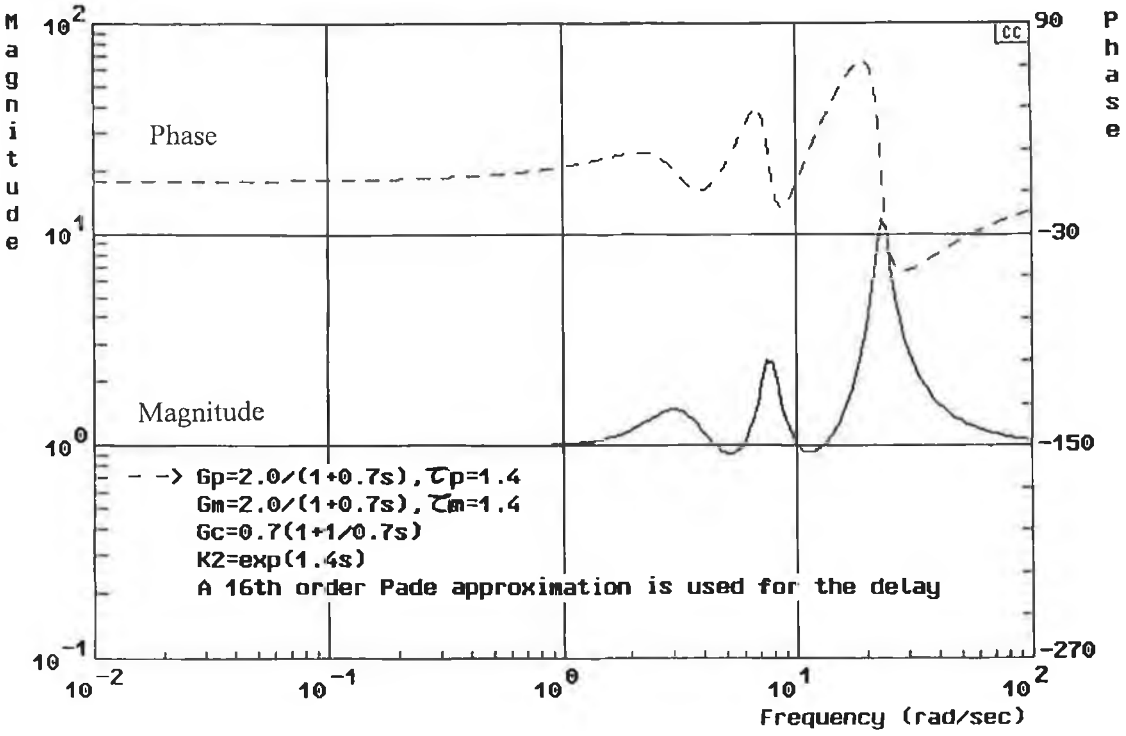


Figure 6.31: $S_{\alpha}^{T^{MSP}} / S_{\alpha}^{T^{SP}}$ when the parameters defined in equations (6.38) to (6.40) are used, with the disturbance present on the process output; the exact value of the time advance is used.



6.5 Conclusions

A modification to the conventional Smith predictor structure for the control of a process with time delay has been proposed to facilitate the achievement of a modest improvement in the closed loop system responses. The modification involves approximating a time advance term that may be incorporated in the outer feedback loop of the predictor. It has been shown analytically and in simulation that the method facilitates performance improvement, particularly when the desired servo response is relatively slow. If the desired servo response is faster, then unless the model fits the process well (for example, when the order of the process equals that of the model), it is less likely that the responses of the modified Smith predictor will be better than those of the Smith predictor. Generally, the performance of the modified Smith predictor tends to be less damped than that of the Smith predictor; the phase lead network in the modified Smith predictor needs to be carefully designed so that system instability does not result. However, the sensitivity of the modified Smith predictor to process parameter variations tends to be less than that of the Smith predictor at lower frequencies.

The work has considered a number of possible model and primary controller combinations; a number of other avenues of exploration are

(a) The use of a higher order term for B(s), when the model is of higher order. There may be some advantage in using a second (or higher) order term for B(s), as is shown by some trial and error supplementary simulation results. However, even in these simulations, the improvement appears to be relatively marginal over the responses determined when a first order term is used for B(s).

(b) The use of higher order process models (perhaps with a zero), and corresponding or reduced order controllers. This is unlikely to have a dramatic effect on the performance of the modified Smith predictor, though it is of course desirable that the mismatch between the model and the nominal process be as small as possible.

(c) The use of the modified Smith predictor strategy, in conjunction with an adaptive identification strategy. The stability of the modified Smith predictor would need to be guaranteed during the updating of the parameters. This may mean that the servo response achieved will be relatively slow, at least until the model and process parameters are closely matched. An adaptive identification procedure to estimate the process parameters in the modified Smith predictor structure is discussed in Chapter 7.

(d) The formulation of robust stability and performance criteria for the modified Smith predictor, similar to those formulated for the Smith predictor by Morari and Zafiriou (1989). Some preliminary work shows that, unsurprisingly, the robust design of the modified Smith predictor is relatively complex. It appears inevitable that an iterative approach to determining the correct robust compensator will be required. This topic is considered in more detail by O'Dwyer (1996f).

(e) Finally, the modified Smith predictor detailed has been chosen to improve the regulator response (when compared to the corresponding response of the Smith predictor); other modified Smith predictors could be defined, for instance, to reduce the effect of the mismatch between the process and model parameters.

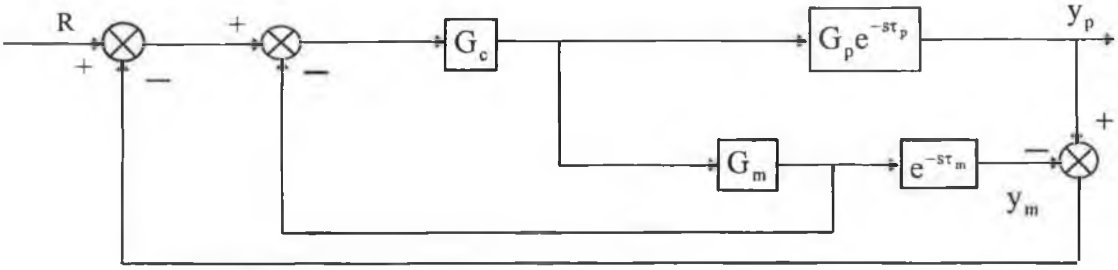
Chapter 7

Closed loop time domain gradient methods for parameter and time delay estimation

7.1 Introduction

The estimation of the model parameters (including time delay), using gradient methods, in a Smith predictor structure, is the subject of this chapter. This topic has been considered by Marshall (1979), (1980), Bahill (1983), Malik-Zafarei and Jamshidi (1987) and O'Connor (1989), amongst others (see Chapter 2); the work in this chapter will expand on the ideas and algorithms considered by these authors. Fundamentally, it is desired to reduce the mismatch between the (unknown) process and the model (particularly the time delay mismatch), so that the performance of the Smith predictor may be improved. The Smith predictor is discussed in detail in Chapters 5 and 6. A block diagram of the Smith predictor is shown in Figure 7.1; G_c, G_m and G_p are functions of the Laplace variable, s.

Figure 7.1: Block diagram of the Smith predictor



As detailed in Chapter 1, it is often assumed that an adequate model for the process is of FOLPD structure, implying that three parameters (gain, time constant and time delay) need to be estimated. It will be shown that the modelling of the process by a more general structure involves a straightforward generalisation of the procedure used to estimate appropriate FOLPD model parameters. The model parameters are updated, in closed loop, based on minimising the integral of the square of the error function (with the error function being the process output minus the model output), using an appropriate gradient algorithm.

Six algorithms (three based on a Gauss-Newton gradient approach and three based on a Newton-Raphson gradient approach) will be implemented for the estimation of the model parameters. The first of the Gauss-Newton based algorithms is outlined by Marshall (1979), (1980) and Bahill (1983); the other two algorithms (which are developed by the author) are refinements on the above algorithm, which eliminate some of the assumptions used in its development, in an attempt to increase the applicability of the procedure. The three Newton-Raphson based algorithms developed by the author correspond directly to the three Gauss-Newton based algorithms; these former algorithms eliminate other assumptions in the development of the estimation procedure, and therefore should increase its applicability still further.

All of these algorithms will also be applied to the estimation of the process parameters in a modified Smith predictor structure (the modified Smith predictor is used to facilitate an improvement in the regulator response of the compensated system in Chapter 6). The convergence properties of the algorithms will also be investigated analytically.

7.2 Algorithms based on a Gauss-Newton gradient approach

7.2.1 Theoretical development of the Gauss-Newton (1) algorithm

Marshall (1979), (1980) and Bahill (1983) have developed a parameter identification algorithm to estimate the corresponding model parameters and time delay of a FOLPD process, in a Smith predictor structure. In the development of the algorithm, the authors assume that the plant output is linearly related to any changes in the plant parameters (assumption 1) i.e.

$$y_p(t,\alpha+\Delta\alpha) = y_p(t,\alpha) + \Delta\alpha\frac{\partial y}{\partial\alpha} \tag{7.1}$$

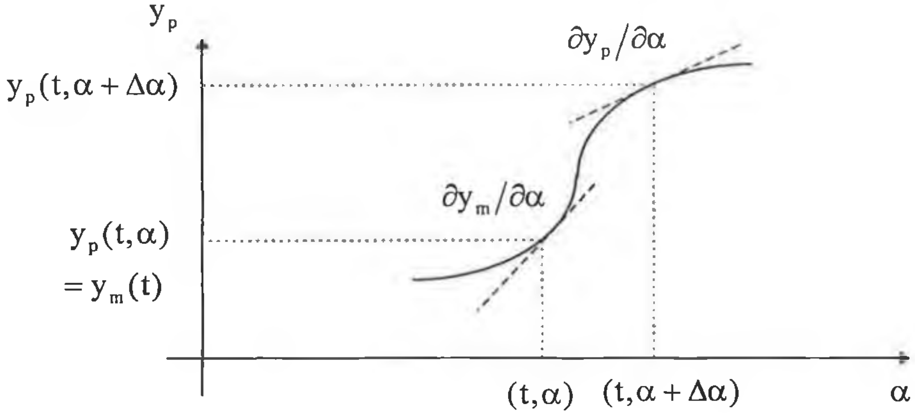
with $y_p(t,\alpha+\Delta\alpha) =$ process output after a change $\Delta\alpha$ in parameter α,

$y_p(t,\alpha) =$ starting value of the process output = model output $y_m(t)$ and

$e(t,\alpha+\Delta\alpha) = y_p(t,\alpha+\Delta\alpha) - y_p(t,\alpha) = \Delta\alpha\frac{\partial y}{\partial\alpha}$.

This assumption effectively means that the change in the parameter being updated is assumed to be small. The idea may be represented graphically as shown in Figure 7.2 .

Figure 7.2: Graphical interpretation of the algorithm



$e(t,\alpha+\Delta\alpha)$ may be estimated in a number of different ways; three such methods, based on Figure 7.2, are as follows:

$$e(t,\alpha+\Delta\alpha) \approx \Delta\alpha\frac{\partial y_p}{\partial\alpha} \tag{7.2}$$

$$e(t,\alpha+\Delta\alpha) \approx \Delta\alpha\frac{\partial y_m}{\partial\alpha} \tag{7.3}$$

$$e(t,\alpha+\Delta\alpha) \approx 0.5\Delta\alpha\left(\frac{\partial y_p}{\partial\alpha}+\frac{\partial y_m}{\partial\alpha}\right) \tag{7.4}$$

Marshall (1979), (1980) and Bahill (1983) use equation (7.2) in their development of the identification method (assumption 2), for updating the gain, time constant and time delay of a FOLPD process model. This development is provided in detail; for convenience in the development, the dependent variables associated with the parameters are henceforth not shown explicitly. From equations (7.1) and (7.2)

$$y_p \approx y_m + \Delta\alpha\frac{\partial y_p}{\partial\alpha} \text{ i.e. } e \approx \frac{\partial y_p}{\partial K_p}\Delta K_m + \frac{\partial y_p}{\partial T_p}\Delta T_m + \frac{\partial y_p}{\partial \tau_p}\Delta\tau_m \tag{7.5}$$

with $\Delta K_m, \Delta T_m$ and $\Delta\tau_m$ being the desired change in the model gain, time constant and time delay, respectively. It is desired to update the model parameters to minimise the following cost function:

$$J = 0.5\int e^2 dt \tag{7.6}$$

From equations (7.5) and (7.6), the cost function is expressed as

$$J = 0.5\int\left(\frac{\partial y_p}{\partial K_p}\Delta K_m + \frac{\partial y_p}{\partial T_p}\Delta T_m + \frac{\partial y_p}{\partial \tau_p}\Delta\tau_m\right)^2 dt \tag{7.7}$$

If it is assumed that $\Delta K_m, \Delta T_m$ and $\Delta\tau_m$ are time invarient (assumption 3), then, from equation (7.7),

$$J = 0.5\Delta K_m{}^2\int\left(\frac{\partial y_p}{\partial K_p}\right)^2 dt + 0.5\Delta T_m{}^2\int\left(\frac{\partial y_p}{\partial T_p}\right)^2 dt + 0.5\Delta\tau_m{}^2\int\left(\frac{\partial y_p}{\partial \tau_p}\right)^2 dt$$

$$+\Delta\tau_m\Delta K_m\int\frac{\partial y_p}{\partial\tau_p}\frac{\partial y_p}{\partial K_p}dt+\Delta\tau_m\Delta T_m\int\frac{\partial y_p}{\partial\tau_p}\frac{\partial y_p}{\partial T_p}dt+\Delta K_m\Delta T_m\int\frac{\partial y_p}{\partial K_p}\frac{\partial y_p}{\partial T_p}dt \quad (7.8)$$

Therefore, from equation (7.8) and the application of assumption 1 (O'Dwyer (1996i)), it may be shown that

$$\frac{\partial J}{\partial K_p}\approx\Delta K_m\int\left(\frac{\partial y_p}{\partial K_p}\right)^2dt+\Delta\tau_m\int\frac{\partial y_p}{\partial\tau_p}\frac{\partial y_p}{\partial K_p}dt+\Delta T_m\int\frac{\partial y_p}{\partial K_p}\frac{\partial y_p}{\partial T_p}dt \quad (7.9)$$

$$\frac{\partial J}{\partial T_p}\approx\Delta T_m\int\left(\frac{\partial y_p}{\partial T_p}\right)^2dt+\Delta\tau_m\int\frac{\partial y_p}{\partial\tau_p}\frac{\partial y_p}{\partial T_p}dt+\Delta K_m\int\frac{\partial y_p}{\partial K_p}\frac{\partial y_p}{\partial T_p}dt \quad (7.10)$$

$$\frac{\partial J}{\partial\tau_p}\approx\Delta\tau_m\int\left(\frac{\partial y_p}{\partial\tau_p}\right)^2dt+\Delta K_m\int\frac{\partial y_p}{\partial\tau_p}\frac{\partial y_p}{\partial K_p}dt+\Delta T_m\int\frac{\partial y_p}{\partial\tau_p}\frac{\partial y_p}{\partial T_p}dt \quad (7.11)$$

The development of equations (7.9) to (7.11) is equivalent to the development of the following approximation:

$$\frac{\partial^2 J}{\partial\alpha_1\partial\alpha_2}\approx\int e\frac{\partial^2 y_p}{\partial\alpha_1\partial\alpha_2}+\frac{\partial y_p}{\partial\alpha_1}\frac{\partial y_p}{\partial\alpha_2}dt\approx\int\frac{\partial y_p}{\partial\alpha_1}\frac{\partial y_p}{\partial\alpha_2}dt \quad (7.12)$$

with $\alpha_1,\alpha_2=[K_p,T_p,\tau_p]^T$ i.e. the algorithm defined by equations (7.9) to (7.11) is a gradient algorithm of Gauss-Newton form (as developed in Chapter 2, Section 2.2.4.1). Now, using

$$\frac{\partial J}{\partial\alpha}=\frac{\partial}{\partial\alpha}\left[0.5\int e^2dt\right] \quad (7.13)$$

and if it is assumed that (assumption 4)

$$\frac{\partial}{\partial\alpha}\left[0.5\int e^2dt\right]=0.5\int\frac{\partial}{\partial\alpha}\left[e^2\right]dt \quad (7.14)$$

therefore
$$\frac{\partial J}{\partial \alpha} = \int e \frac{\partial e}{\partial \alpha} dt \qquad (7.15)$$

with $\alpha = [K_p, T_p, \tau_p]^T$. A further assumption made by the authors is that

$$\frac{\partial e}{\partial \alpha} = \frac{\partial y_p}{\partial \alpha} - \frac{\partial y_m}{\partial \alpha} \approx \frac{\partial y_p}{\partial \alpha} \qquad (7.16)$$

i.e. $\partial y_m / \partial \alpha << \partial y_p / \partial \alpha$. Further analysis shows that this assumption is equivalent to the application of assumption 1 and assumption 2 (O'Dwyer (1996i)). The update values of the model gain, time constant and time delay are calculated, using equations (7.9), (7.10), (7.11), (7.15) and (7.16), by solving the following relationship:

$$\begin{bmatrix} \int e \frac{\partial y_p}{\partial K_p} dt \\ \int e \frac{\partial y_p}{\partial T_p} dt \\ \int e \frac{\partial y_p}{\partial \tau_p} dt \end{bmatrix} = \begin{bmatrix} \int \left(\frac{\partial y_p}{\partial K_p}\right)^2 dt & \int \left(\frac{\partial y_p}{\partial T_p}\right)\left(\frac{\partial y_p}{\partial K_p}\right) dt & \int \left(\frac{\partial y_p}{\partial \tau_p}\right)\left(\frac{\partial y_p}{\partial K_p}\right) dt \\ \int \left(\frac{\partial y_p}{\partial T_p}\right)\left(\frac{\partial y_p}{\partial K_p}\right) dt & \int \left(\frac{\partial y_p}{\partial T_p}\right)^2 dt & \int \left(\frac{\partial y_p}{\partial T_p}\right)\left(\frac{\partial y_p}{\partial \tau_p}\right) dt \\ \int \left(\frac{\partial y_p}{\partial \tau_p}\right)\left(\frac{\partial y_p}{\partial K_p}\right) dt & \int \left(\frac{\partial y_p}{\partial T_p}\right)\left(\frac{\partial y_p}{\partial \tau_p}\right) dt & \int \left(\frac{\partial y_p}{\partial \tau_p}\right)^2 dt \end{bmatrix} \begin{bmatrix} \Delta K_m \\ \Delta T_m \\ \Delta \tau_m \end{bmatrix} \qquad (7.17)$$

Bahill (1983) develops the procedure described for updating the time delay only; Marshall (1979), (1980) develops the procedure to facilitate updating of all of the parameters. Of course, the updating of the gain term only and the time constant term only could also be implemented using the algorithm.

The implementation of the procedure requires the evaluation of $\partial y_p / \partial \alpha$. The response of the Smith predictor (Figure 7.1) is calculated to be

$$\frac{y_p}{R} = \frac{G_c G_p e^{-s\tau_p}}{1 + G_c G_m + G_c \left[G_p e^{-s\tau_p} - G_m e^{-s\tau_m} \right]} \qquad (7.18)$$

The following formulae have been developed from equation (7.18) for the evaluation of the required partial derivatives (or "sensitivity functions") in equation (7.17):

$$\frac{\partial y_p}{\partial \tau_p} = R\frac{-\left(\dfrac{sG_cK_pe^{-s\tau_p}}{1+sT_p}\right)\left(1+\dfrac{G_cK_m}{1+sT_m}-\dfrac{G_cK_me^{-s\tau_m}}{1+sT_m}\right)}{\left(1+\dfrac{G_cK_m}{1+sT_m}-\dfrac{G_cK_me^{-s\tau_m}}{1+sT_m}+\dfrac{G_cK_pe^{-s\tau_p}}{1+sT_p}\right)^2} = -sy_p\left(1-\frac{y_p}{R}\right) \tag{7.19}$$

$$\frac{\partial y_p}{\partial K_p} = R\frac{\left(\dfrac{G_ce^{-s\tau_p}}{1+sT_p}\right)\left(1+\dfrac{G_cK_m}{1+sT_m}-\dfrac{G_cK_me^{-s\tau_m}}{1+sT_m}\right)}{\left(1+\dfrac{G_cK_m}{1+sT_m}-\dfrac{G_cK_me^{-s\tau_m}}{1+sT_m}+\dfrac{G_cK_pe^{-s\tau_p}}{1+sT_p}\right)^2} = \frac{1}{K_p}y_p\left(1-\frac{y_p}{R}\right) \tag{7.20}$$

$$\frac{\partial y_p}{\partial T_p} = R\frac{\left(-\dfrac{sG_cK_pe^{-s\tau_p}}{(1+sT_p)^2}\right)\left(1+\dfrac{G_cK_m}{1+sT_m}-\dfrac{G_cK_me^{-s\tau_m}}{1+sT_m}\right)}{\left(1+\dfrac{G_cK_m}{1+sT_m}-\dfrac{G_cK_me^{-s\tau_m}}{1+sT_m}+\dfrac{G_cK_pe^{-s\tau_p}}{1+sT_p}\right)^2} = -\frac{s}{\left(1+sT_p\right)}y_p\left(1-\frac{y_p}{R}\right) \tag{7.21}$$

Alternative Gauss-Newton gradient algorithms, based on relaxing the assumptions used in the procedure, may be derived. Two such algorithms are detailed in Sections 7.2.2 and 7.2.3.

7.2.2 Theoretical development of the Gauss-Newton (2) algorithm

The final version of the Gauss-Newton (1) algorithm (equation (7.17)) involved assuming that $\partial y_p/\partial\alpha \approx \partial e/\partial\alpha$ (equation (7.16)). However, it is straightforward to calculate $\partial y_m/\partial\alpha$, and thus the assumption in equation (7.16) may be eliminated. The update values of the model gain, time constant and time delay are subsequently calculated by solving the following relationship (which corresponds to equation (7.17)):

$$\begin{bmatrix} \int e\dfrac{\partial e}{\partial K_p}dt \\ \int e\dfrac{\partial e}{\partial T_p}dt \\ \int e\dfrac{\partial e}{\partial \tau_p}dt \end{bmatrix} = \begin{bmatrix} \int\left(\dfrac{\partial y_p}{\partial K_p}\right)^2 dt & \int\left(\dfrac{\partial y_p}{\partial T_p}\right)\left(\dfrac{\partial y_p}{\partial K_p}\right)dt & \int\left(\dfrac{\partial y_p}{\partial \tau_p}\right)\left(\dfrac{\partial y_p}{\partial K_p}\right)dt \\ \int\left(\dfrac{\partial y_p}{\partial T_p}\right)\left(\dfrac{\partial y_p}{\partial K_p}\right)dt & \int\left(\dfrac{\partial y_p}{\partial T_p}\right)^2 dt & \int\left(\dfrac{\partial y_p}{\partial T_p}\right)\left(\dfrac{\partial y_p}{\partial \tau_p}\right)dt \\ \int\left(\dfrac{\partial y_p}{\partial \tau_p}\right)\left(\dfrac{\partial y_p}{\partial K_p}\right)dt & \int\left(\dfrac{\partial y_p}{\partial T_p}\right)\left(\dfrac{\partial y_p}{\partial \tau_p}\right)dt & \int\left(\dfrac{\partial y_p}{\partial \tau_p}\right)^2 dt \end{bmatrix} \begin{bmatrix} \Delta K_m \\ \Delta T_m \\ \Delta \tau_m \end{bmatrix} \qquad (7.22)$$

If, in addition, it is assumed that

$$y_p(t,\alpha+\Delta\alpha) = y_p(t,\alpha) + \Delta\alpha\frac{\partial e}{\partial\alpha} \qquad (7.23)$$

then the update values of the model parameters are calculated by solving the following relationship (which is obtained using a similar procedure to that given in equations (7.5) to (7.11)):

$$\begin{bmatrix} \int e\dfrac{\partial e}{\partial K_p}dt \\ \int e\dfrac{\partial e}{\partial T_p}dt \\ \int e\dfrac{\partial e}{\partial \tau_p}dt \end{bmatrix} = \begin{bmatrix} \int\left(\dfrac{\partial e}{\partial K_p}\right)^2 dt & \int\left(\dfrac{\partial e}{\partial T_p}\right)\left(\dfrac{\partial e}{\partial K_p}\right)dt & \int\left(\dfrac{\partial e}{\partial \tau_p}\right)\left(\dfrac{\partial e}{\partial K_p}\right)dt \\ \int\left(\dfrac{\partial e}{\partial T_p}\right)\left(\dfrac{\partial e}{\partial K_p}\right)dt & \int\left(\dfrac{\partial e}{\partial T_p}\right)^2 dt & \int\left(\dfrac{\partial e}{\partial T_p}\right)\left(\dfrac{\partial e}{\partial \tau_p}\right)dt \\ \int\left(\dfrac{\partial e}{\partial \tau_p}\right)\left(\dfrac{\partial e}{\partial K_p}\right)dt & \int\left(\dfrac{\partial e}{\partial T_p}\right)\left(\dfrac{\partial e}{\partial \tau_p}\right)dt & \int\left(\dfrac{\partial e}{\partial \tau_p}\right)^2 dt \end{bmatrix} \begin{bmatrix} \Delta K_m \\ \Delta T_m \\ \Delta \tau_m \end{bmatrix} \qquad (7.24)$$

This latter assumption is non-optimal; this topic is discussed in more detail in the conclusions of the chapter.

The implementation of the procedure requires the evaluation of $\partial e/\partial\alpha$, $e = y_p - y_m$. The equation for y_m is derived from Figure 7.1 to be

$$y_m = \frac{G_c G_m e^{-s\tau_m} R}{1 + G_c G_m + G_c\left[G_p e^{-s\tau_p} - G_m e^{-s\tau_m}\right]} \qquad (7.25)$$

with y_p given by equation (7.18).

The following formulae have subsequently been developed (from equations (7.18) and (7.25)) for the evaluation of the required partial derivatives (or sensitivity functions) in equation (7.24):

$$\frac{\partial e}{\partial \tau_p} = R\frac{\left(-\frac{sG_cK_pe^{-s\tau_p}}{1+sT_p}\right)\left(1+\frac{G_cK_m}{1+sT_m}\right)}{\left(1+\frac{G_cK_m}{1+sT_m}-\frac{G_cK_me^{-s\tau_m}}{1+sT_m}+\frac{G_cK_pe^{-s\tau_p}}{1+sT_p}\right)^2} = -sy_p\left(1-\frac{y_p}{R}+\frac{y_m}{R}\right) \qquad (7.26)$$

$$\frac{\partial e}{\partial K_p} = R\frac{\left(\frac{G_ce^{-s\tau_p}}{1+sT_p}\right)\left(1+\frac{G_cK_m}{1+sT_m}\right)}{\left(1+\frac{G_cK_m}{1+sT_m}-\frac{G_cK_me^{-s\tau_m}}{1+sT_m}+\frac{G_cK_pe^{-s\tau_p}}{1+sT_p}\right)^2} = \frac{1}{K_p}y_p\left(1-\frac{y_p}{R}+\frac{y_m}{R}\right) \qquad (7.27)$$

$$\frac{\partial e}{\partial T_p} = R\frac{\left(-\frac{sG_cK_pe^{-s\tau_p}}{(1+sT_p)^2}\right)\left(1+\frac{G_cK_m}{1+sT_m}\right)}{\left(1+\frac{G_cK_m}{1+sT_m}-\frac{G_cK_me^{-s\tau_m}}{1+sT_m}+\frac{G_cK_pe^{-s\tau_p}}{1+sT_p}\right)^2} = -\frac{s}{\left(1+sT_p\right)}y_p\left(1-\frac{y_p}{R}+\frac{y_m}{R}\right) \qquad (7.28)$$

7.2.3 Theoretical development of the Gauss-Newton (3) algorithm

The Gauss-Newton (1) algorithm developed in Section 7.2.1 assumed that $e(t,\alpha+\Delta\alpha) \approx \Delta\alpha\,\partial y_p/\partial\alpha$ (equation (7.2)). As was mentioned in Section 7.2.1, an alternative approximation is given by equation (7.4). This approximation and other complementary assumptions to those made by Bahill (1983) are used in the development of the method in this section. It will also be assumed that

$$\frac{\partial e}{\partial \alpha} \approx 0.5\left(\frac{\partial y_p}{\partial \alpha}+\frac{\partial y_m}{\partial \alpha}\right) \qquad (7.29)$$

As in Section 7.2.2, this assumption is non-optimal; this topic is discussed in more detail in the conclusions of the chapter.

The update values of the model parameters are subsequently calculated by solving the following relationship (which is obtained using a similar procedure to that given in equations (7.5) to (7.16)):

$$\begin{bmatrix} \int e\dfrac{\partial e_1}{\partial K_p}dt \\ \int e\dfrac{\partial e_1}{\partial T_p}dt \\ \int e\dfrac{\partial e_1}{\partial \tau_p}dt \end{bmatrix} = \begin{bmatrix} \int\left(\dfrac{\partial e_1}{\partial K_p}\right)^2 dt & \int\left(\dfrac{\partial e_1}{\partial T_p}\right)\left(\dfrac{\partial e_1}{\partial K_p}\right)dt & \int\left(\dfrac{\partial e_1}{\partial \tau_p}\right)\left(\dfrac{\partial e_1}{\partial K_p}\right)dt \\ \int\left(\dfrac{\partial e_1}{\partial T_p}\right)\left(\dfrac{\partial e_1}{\partial K_p}\right)dt & \int\left(\dfrac{\partial e_1}{\partial T_p}\right)^2 dt & \int\left(\dfrac{\partial e_1}{\partial T_p}\right)\left(\dfrac{\partial e_1}{\partial \tau_p}\right)dt \\ \int\left(\dfrac{\partial e_1}{\partial \tau_p}\right)\left(\dfrac{\partial e_1}{\partial K_p}\right)dt & \int\left(\dfrac{\partial e_1}{\partial T_p}\right)\left(\dfrac{\partial e_1}{\partial \tau_p}\right)dt & \int\left(\dfrac{\partial e_1}{\partial \tau_p}\right)^2 dt \end{bmatrix} \begin{bmatrix} \Delta K_m \\ \Delta T_m \\ \Delta \tau_m \end{bmatrix} \qquad (7.30)$$

with $e_1 = y_p + y_m$. The following formulae have been developed for the evaluation of the required partial derivatives (or sensitivity functions) in equation (7.30), based on equations (7.18) and (7.25):

$$\frac{\partial(y_p + y_m)}{\partial \tau_p} = R\frac{\left(-\dfrac{sG_cK_pe^{-s\tau_p}}{1+sT_p}\right)\left(1+\dfrac{G_cK_m}{1+sT_m}-\dfrac{2G_cK_me^{-s\tau_m}}{1+sT_m}\right)}{\left(1+\dfrac{G_cK_m}{1+sT_m}-\dfrac{G_cK_me^{-s\tau_m}}{1+sT_m}+\dfrac{G_cK_pe^{-s\tau_p}}{1+sT_p}\right)^2} = -sy_p\left(1-\frac{y_p}{R}-\frac{y_m}{R}\right)$$

(7.31)

$$\frac{\partial(y_p + y_m)}{\partial K_p} = R\frac{\left(\dfrac{G_ce^{-s\tau_p}}{1+sT_p}\right)\left(1+\dfrac{G_cK_m}{1+sT_m}-\dfrac{2G_cK_me^{-s\tau_m}}{1+sT_m}\right)}{\left(1+\dfrac{G_cK_m}{1+sT_m}-\dfrac{G_cK_me^{-s\tau_m}}{1+sT_m}+\dfrac{G_cK_pe^{-s\tau_p}}{1+sT_p}\right)^2} = \frac{1}{K_p}y_p\left(1-\frac{y_p}{R}-\frac{y_m}{R}\right)$$

(7.32)

$$\frac{\partial(y_p + y_m)}{\partial T_p} = R\frac{\left(-\dfrac{sG_cK_pe^{-s\tau_p}}{(1+sT_p)^2}\right)\left(1+\dfrac{G_cK_m}{1+sT_m}-\dfrac{2G_cK_me^{-s\tau_m}}{1+sT_m}\right)}{\left(1+\dfrac{G_cK_m}{1+sT_m}-\dfrac{G_cK_me^{-s\tau_m}}{1+sT_m}+\dfrac{G_cK_pe^{-s\tau_p}}{1+sT_p}\right)^2} = -\frac{s}{\left(1+sT_p\right)}y_p\left(1-\frac{y_p}{R}-\frac{y_m}{R}\right)$$

(7.33)

7.2.4 Algorithm representations

The block diagram representation of the algorithms for updating the model time delay only is provided in Figure 7.3. The block diagram of the three algorithms used differs only in the sensitivity function; $v = y_p$ for the Gauss-Newton (1) algorithm, $v = e$ for the Gauss-Newton (2) algorithm and $v = e_1$ for the Gauss-Newton (3) algorithm ('x' = multiply).

The block diagram representation of the algorithms for updating the model gain or time constant only, assuming a FOLPD model, is provided in Figure 7.4. In the sensitivity function in the diagram, $\omega = K_p$ when the model gain is to be updated and $\omega = T_p$ when the model time constant is to be updated. The code for v is identical to that in Figure 7.3.

The block diagram representation of the sensitivity functions, for each of the Gauss-Newton algorithms, is provided in Figure 7.5; K_m', T_m' and τ_m' are approximations for the (unknown) process parameter values used in the sensitivity functions (the explicit dependence of the sensitivity functions on the process parameter values is shown in equations (7.19) to (7.21), for example). In this diagram, the sensitivity functions $\partial y_p/\partial K_p$, $\partial y_p/\partial T_p$ and $\partial y_p/\partial \tau_p$ correspond to the Gauss-Newton (1) algorithm, the sensitivity functions $\partial e/\partial K_p$, $\partial e/\partial T_p$ and $\partial e/\partial \tau_p$ correspond to the Gauss-Newton (2) algorithm and the sensitivity functions $\partial e_1/\partial K_p$, $\partial e_1/\partial T_p$ and $\partial e_1/\partial \tau_p$ correspond to the Gauss-Newton (3) algorithm.

The block diagram representation of the algorithms to facilitate the updating of all of the parameters simultaneously is provided in Figure 7.6. The code for v is identical to that in Figure 7.3.

There is a straightforward extension of the algorithms to allow updating of the parameters of an arbitrary order model plus time delay; the block diagram representations of the relevant algorithms are provided in Figures 7.7 and 7.8.

A simplified version of the algorithms for updating all of the parameters simultaneously may be implemented by approximating the matrices given by equations (7.17), (7.24) or (7.30) by a diagonal matrix, with the same diagonal entries as in the defined matrices. The appropriateness of using such an approach has been evaluated by determining the ratio of the magnitude of the off-diagonal elements to the diagonal elements; however, the ratios calculated depend on the (unknown) process parameter values and are also frequency dependent. In practice, therefore, the appropriateness of using these simplified algorithms would have to be evaluated in simulation.

Some preliminary comments are appropriate on the algorithms:

(a) The calculation of the sensitivity functions $\partial v/\partial \tau_p$, $v = y_p$, e or e_1 (Figures 7.5 and 7.8) involve the use of a derivative term, which may be problematic in the presence of noise. The use of a filtered derivative term should be helpful.

(b) It is suggested by Bahill (1983) that the values of K_m', T_m' and τ_m' should be put equal to the relevant model parameters, with these values updated "as often as possible".

(c) The assumptions made in deriving the methods limit their application to the identification of processes whose parameters "change slowly" (Bahill (1983)).

Figure 7.3: Representation of the Gauss-Newton algorithms for time delay estimation.

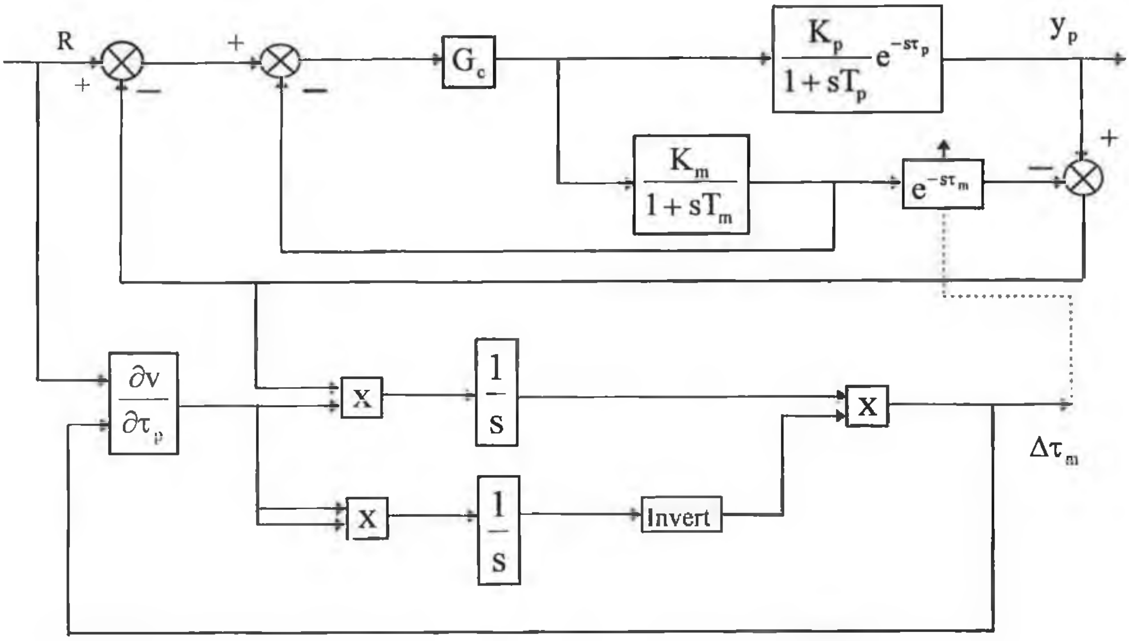


Figure 7.4: Representation of the Gauss-Newton algorithms, for model gain or model time constant estimation.

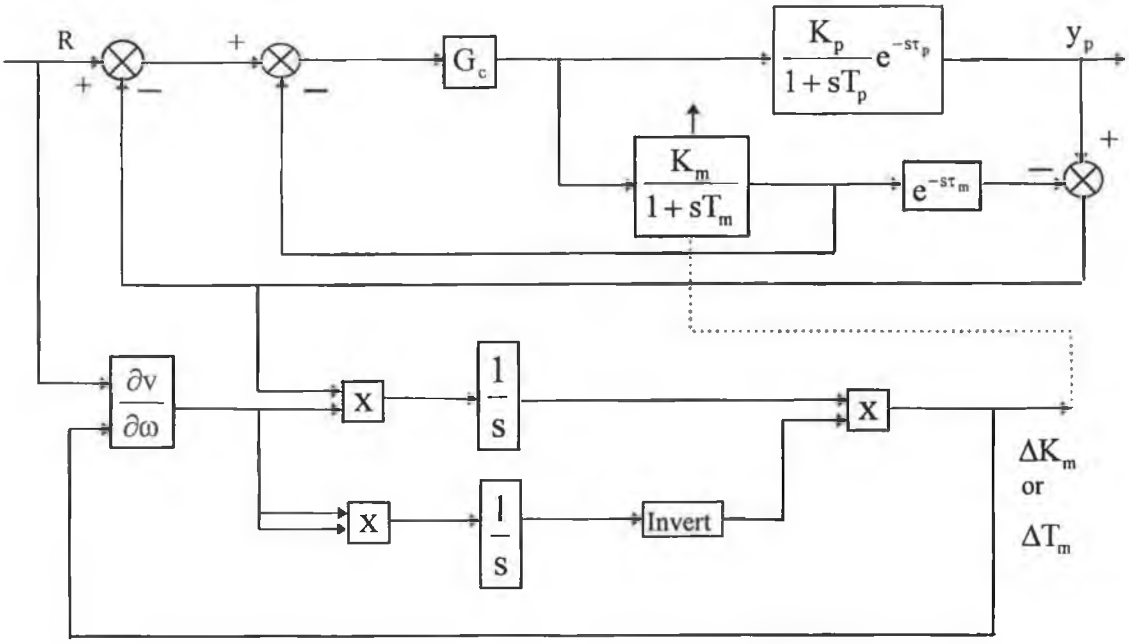


Figure 7.5: Representation of the sensitivity functions for the Gauss-Newton algorithms.

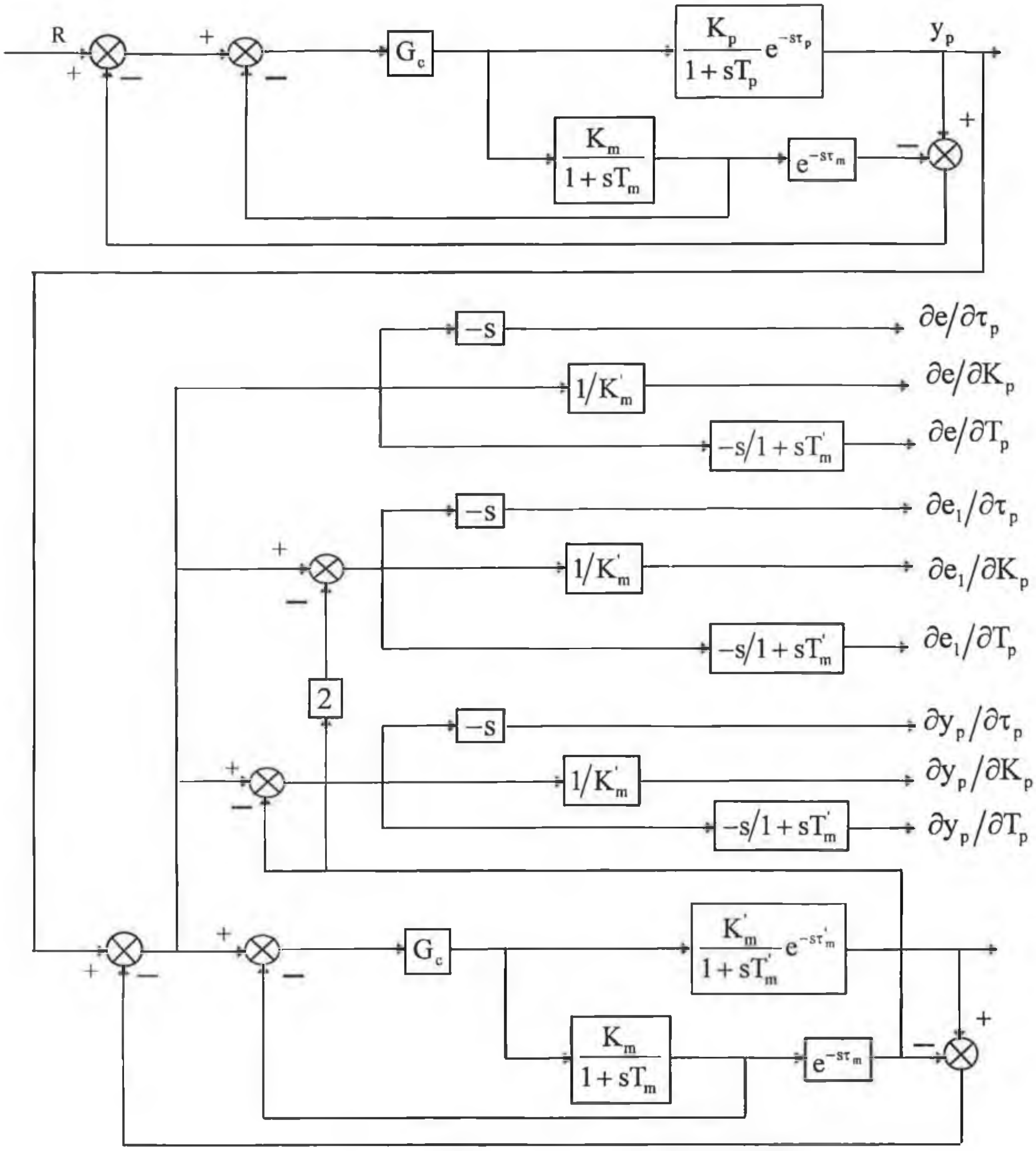
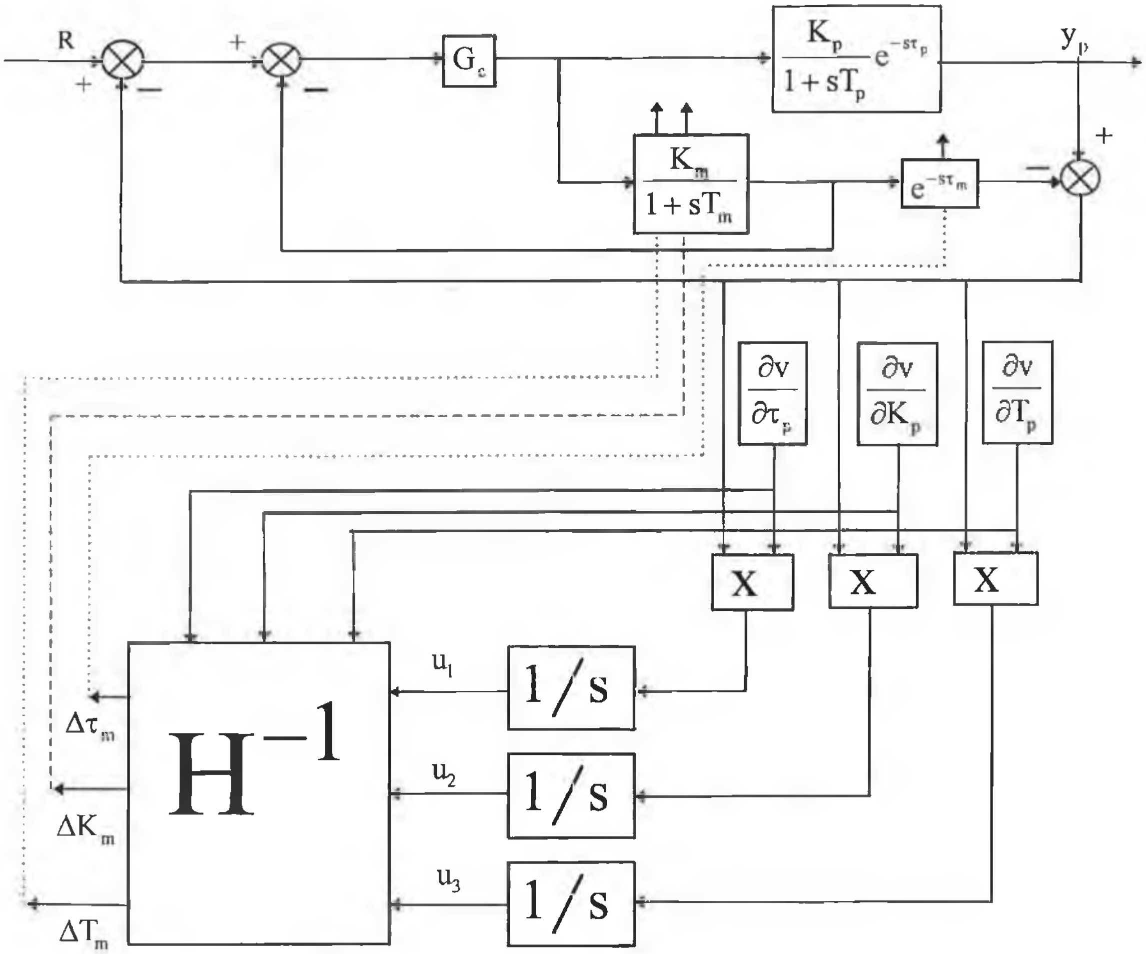


Figure 7.6: Representation of the Gauss-Newton algorithms for simultaneous model parameter estimation.

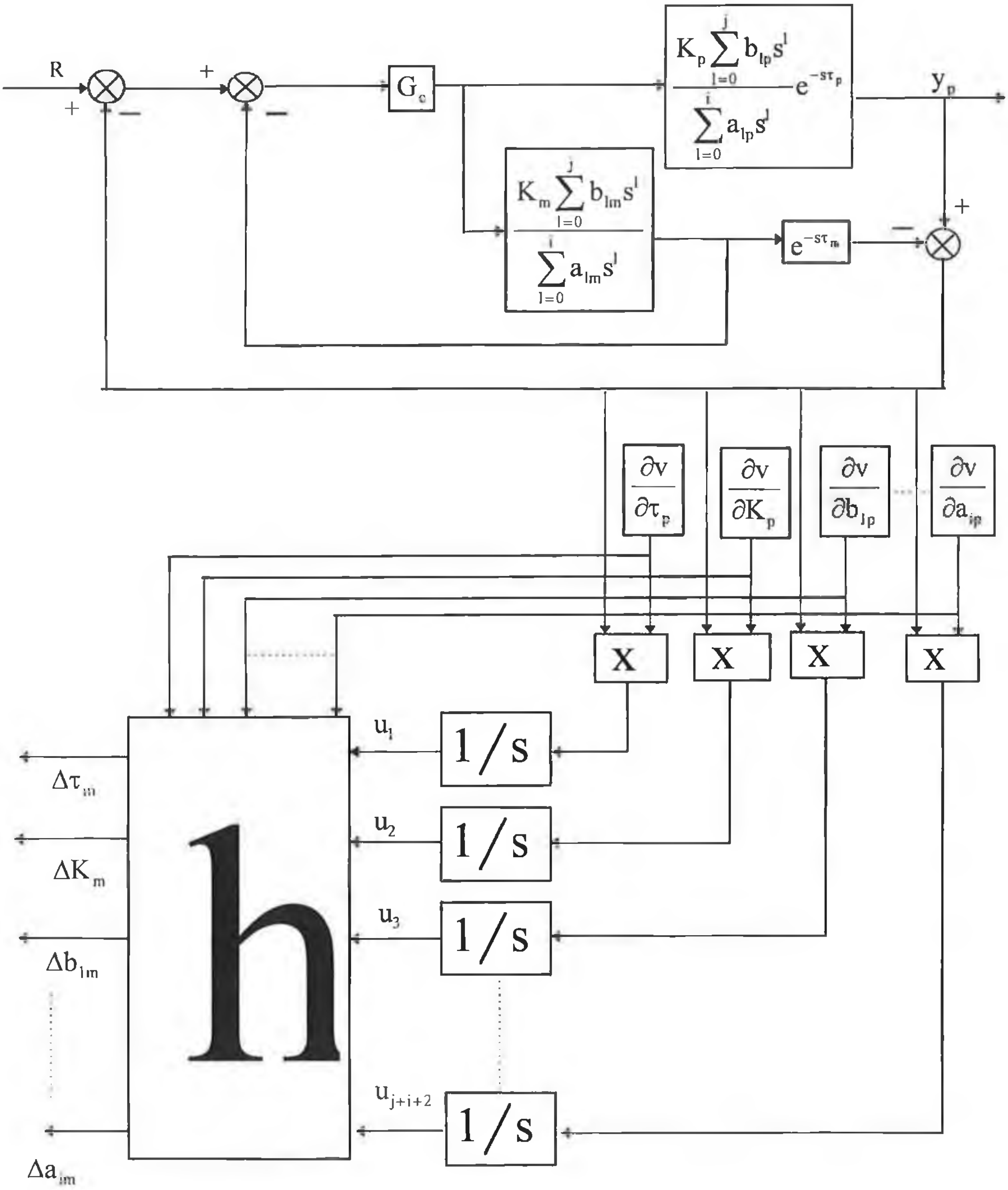


with

$$H=\begin{bmatrix} \int\left(\frac{\partial v}{\partial \tau_p}\right)^2 dt & \int\left(\frac{\partial v}{\partial \tau_p}\right)\left(\frac{\partial v}{\partial K_p}\right)dt & \int\left(\frac{\partial v}{\partial \tau_p}\right)\left(\frac{\partial v}{\partial T_p}\right)dt \\ \int\left(\frac{\partial v}{\partial \tau_p}\right)\left(\frac{\partial v}{\partial K_p}\right)dt & \int\left(\frac{\partial v}{\partial K_p}\right)^2 dt & \int\left(\frac{\partial v}{\partial T_p}\right)\left(\frac{\partial v}{\partial K_p}\right)dt \\ \int\left(\frac{\partial v}{\partial \tau_p}\right)\left(\frac{\partial v}{\partial T_p}\right)dt & \int\left(\frac{\partial v}{\partial T_p}\right)\left(\frac{\partial v}{\partial K_p}\right)dt & \int\left(\frac{\partial v}{\partial T_p}\right)^2 dt \end{bmatrix}$$

and $v = y_p, e$ or e_1, as appropriate. The sensitivity functions are calculated as in Figure 7.5.

Figure 7.7: Representation of the Gauss-Newton algorithms, for simultaneous model parameter estimation (of a general order model).

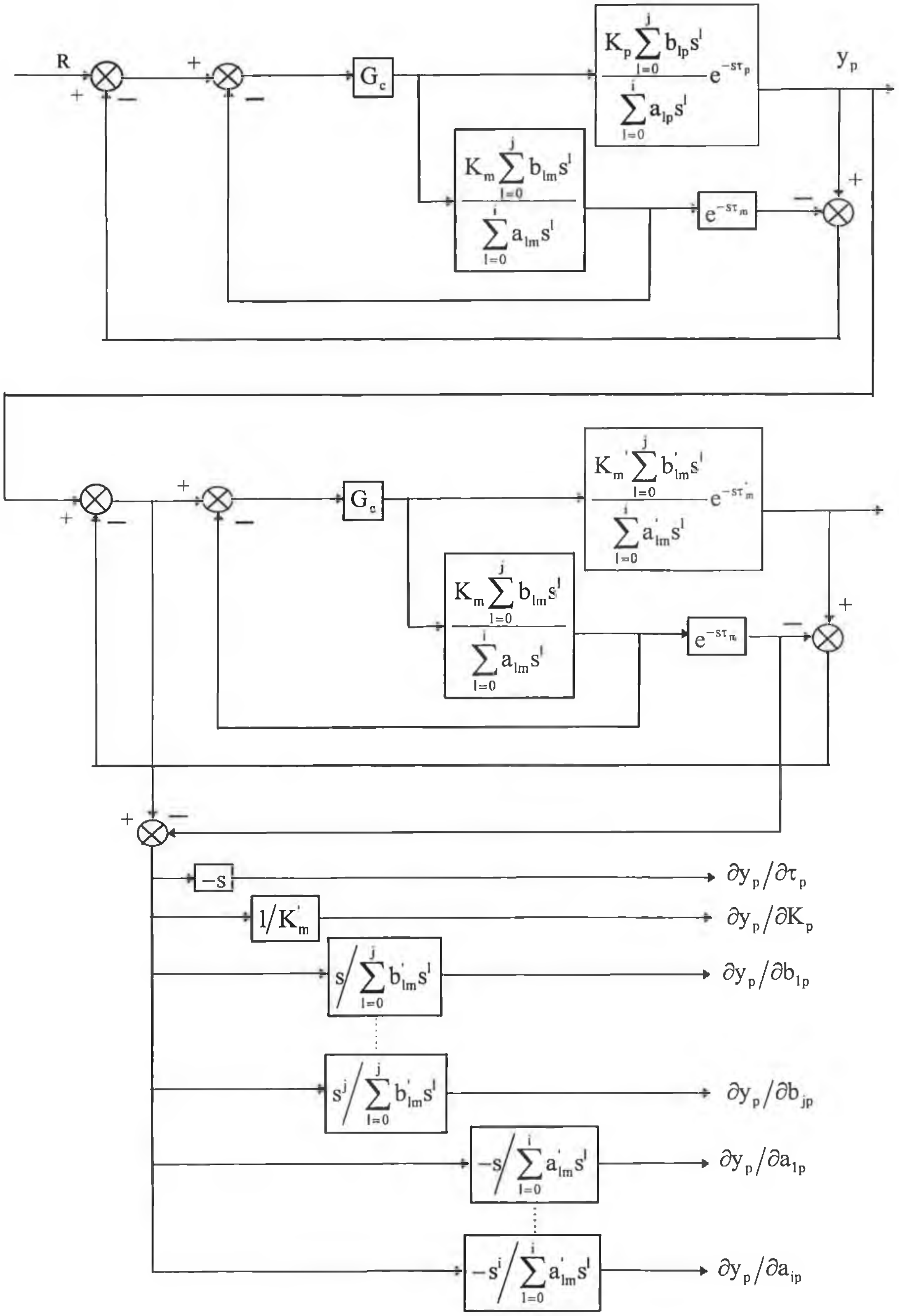


with

$$h = \begin{bmatrix} \int\left(\frac{\partial v}{\partial \tau_p}\right)^2 dt & \int\left(\frac{\partial v}{\partial \tau_p}\right)\left(\frac{\partial v}{\partial K_p}\right)dt & \int\left(\frac{\partial v}{\partial \tau_p}\right)\left(\frac{\partial v}{\partial b_{1p}}\right)dt. & .. & \int\left(\frac{\partial v}{\partial a_{ip}}\right)\left(\frac{\partial v}{\partial \tau_p}\right)dt \\ \int\left(\frac{\partial v}{\partial \tau_p}\right)\left(\frac{\partial v}{\partial K_p}\right)dt & \int\left(\frac{\partial v}{\partial K_p}\right)^2 dt & \int\left(\frac{\partial v}{\partial b_{1p}}\right)\left(\frac{\partial v}{\partial K_p}\right)dt. & . & \int\left(\frac{\partial v}{\partial a_{ip}}\right)\left(\frac{\partial v}{\partial K_p}\right)dt \\ \int\left(\frac{\partial v}{\partial \tau_p}\right)\left(\frac{\partial v}{\partial b_{1p}}\right)dt & \int\left(\frac{\partial v}{\partial b_{1p}}\right)\left(\frac{\partial v}{\partial K_p}\right)dt & \int\left(\frac{\partial v}{\partial b_{1p}}\right)^2 dt. & . & \int\left(\frac{\partial v}{\partial a_{ip}}\right)\left(\frac{\partial v}{\partial b_{1p}}\right)dt \\ & & & & \\ \int\left(\frac{\partial v}{\partial a_{ip}}\right)\left(\frac{\partial v}{\partial \tau_p}\right)dt & \int\left(\frac{\partial v}{\partial a_{ip}}\right)\left(\frac{\partial v}{\partial K_p}\right)dt & \int\left(\frac{\partial v}{\partial a_{ip}}\right)\left(\frac{\partial v}{\partial b_{1p}}\right) & . \; . & \int\left(\frac{\partial v}{\partial a_{ip}}\right)^2 dt \end{bmatrix}^{-1}$$

and with $v = y_p, e$ or e_1, as appropriate. The sensitivity functions for the Gauss-Newton (1) representation are calculated as shown in Figure 7.8. The sensitivity functions for the Gauss-Newton (2) and Gauss-Newton (3) representations follow the same template as those shown in Figure 7.5; the values of $K_m^{'}$, $\tau_m^{'}$, $b_{lm}^{'}$, $l = 1 \ldots j$ and $a_{lm}^{'}$, $l = 1 \ldots i$ used are approximations for the process parameter values.

Figure 7.8: Representation of the sensitivity functions for the Gauss-Newton algorithms (general order model).



7.3 Algorithms based on a Newton-Raphson gradient approach

The algorithms described in Section 7.2 are based on a Gauss-Newton gradient approach. Such algorithms will facilitate linear convergence of the model parameters to the process parameters (Soderstrom and Stoica (1989)). For quadratic convergence of the model parameters to the process parameters, it is necessary to calculate the second partial derivative of the error with respect to the parameter vector (equation (2.6), Chapter 2). Three such Newton-Raphson gradient algorithms are discussed, corresponding to the three Gauss-Newton algorithms discussed in Section 7.2.

7.3.1 Theoretical development of the Newton-Raphson (1) algorithm

The Gauss-Newton (1) algorithm (Section 7.2.1) has used the approximation that

$$\frac{\partial^2 J}{\partial\alpha_1\partial\alpha_2} \approx \int e\frac{\partial^2 y_p}{\partial\alpha_1\partial\alpha_2} + \frac{\partial y_p}{\partial\alpha_1}\frac{\partial y_p}{\partial\alpha_2}dt \approx \int\frac{\partial y_p}{\partial\alpha_1}\frac{\partial y_p}{\partial\alpha_2}dt$$

(equation (7.12)). However, it is straightforward to calculate $\partial^2 y_p/\partial\alpha_1\,\partial\alpha_2$ (by partially differentiating equations (7.19), (7.20) and (7.21)), facilitating the elimination of the assumption that this term may be neglected (this means that the corresponding gradient approach may be classified as a Newton-Raphson gradient algorithm). The update values of the model parameters may be subsequently calculated, by solving the following relationship (using a similar procedure to that given in Section 7.2.1):

$$\begin{bmatrix} \int e\frac{\partial y_p}{\partial K_p}dt \\ \int e\frac{\partial y_p}{\partial T_p}dt \\ \int e\frac{\partial y_p}{\partial \tau_p}dt \end{bmatrix} = \begin{bmatrix} \int\left(\frac{\partial y_p}{\partial K_p}\right)^2 + e\frac{\partial^2 y_p}{\partial K_p{}^2}dt & \int\left(\frac{\partial y_p}{\partial T_p}\right)\left(\frac{\partial y_p}{\partial K_p}\right) + e\frac{\partial^2 y_p}{\partial T_p\partial K_p}dt & \int\left(\frac{\partial y_p}{\partial \tau_p}\right)\left(\frac{\partial y_p}{\partial K_p}\right) + e\frac{\partial^2 y_p}{\partial \tau_p\partial K_p}dt \\ \int\left(\frac{\partial y_p}{\partial T_p}\right)\left(\frac{\partial y_p}{\partial K_p}\right) + e\frac{\partial^2 y_p}{\partial T_p\partial K_p}dt & \int\left(\frac{\partial y_p}{\partial T_p}\right)^2 + e\frac{\partial^2 y_p}{\partial T_p{}^2}dt & \int\left(\frac{\partial y_p}{\partial T_p}\right)\left(\frac{\partial y_p}{\partial \tau_p}\right) + e\frac{\partial^2 y_p}{\partial \tau_p\partial T_p}dt \\ \int\left(\frac{\partial y_p}{\partial \tau_p}\right)\left(\frac{\partial y_p}{\partial K_p}\right) + e\frac{\partial^2 y_p}{\partial \tau_p\partial K_p}dt & \int\left(\frac{\partial y_p}{\partial T_p}\right)\left(\frac{\partial y_p}{\partial \tau_p}\right) + e\frac{\partial^2 y_p}{\partial T_p\partial \tau_p}dt & \int\left(\frac{\partial y_p}{\partial \tau_p}\right)^2 + e\frac{\partial^2 y_p}{\partial \tau_p{}^2}dt \end{bmatrix} \begin{bmatrix} \Delta K_m \\ \Delta T_m \\ \Delta\tau_m \end{bmatrix} \quad (7.34)$$

The following formulae have been developed for the evaluation of the required second partial derivatives in equation (7.34):

$$\frac{\partial^2 y_p}{\partial \tau_p^{\,2}} = R\frac{\left(\frac{s^2 G_c K_p e^{-s\tau_p}}{1+sT_p}\right)\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}\right)\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}-\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)}{\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}+\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)^3}$$
$$= s^2 y_p\left(1-\frac{y_p}{R}\right)\left(1-2\frac{y_p}{R}\right) \qquad (7.35)$$

$$\frac{\partial^2 y_p}{\partial K_p^{\,2}} = R\frac{\left(-\frac{2G_c^{\,2} e^{-2s\tau_p}}{(1+sT_p)^2}\right)\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}\right)}{\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}+\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)^3} = -\frac{2}{K_p^{\,2}} y_p \frac{y_p}{R}\left(1-\frac{y_p}{R}\right) \qquad (7.36)$$

$$\frac{\partial^2 y_p}{\partial T_p^{\,2}} = R\frac{\left(\frac{2s^2 G_c K_p e^{-s\tau_p}}{(1+sT_p)^3}\right)\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}\right)^2}{\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}+\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)^3} = \frac{2s^2}{(1+sT_p)^2} y_p\left(1-\frac{y_p}{R}\right)^2 \qquad (7.37)$$

$$\frac{\partial^2 y_p}{\partial K_p \partial \tau_p} = R\frac{\left(-\frac{sG_c e^{-s\tau_p}}{1+sT_p}\right)\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}\right)\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}-\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)}{\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}+\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)^3}$$
$$= -\frac{s}{K_p} y_p\left(1-\frac{y_p}{R}\right)\left(1-2\frac{y_p}{R}\right) \qquad (7.38)$$

$$\frac{\partial^2 y_p}{\partial T_p \partial \tau_p} = R\frac{\left(\frac{s^2 G_c K_p e^{-s\tau_p}}{(1+sT_p)^2}\right)\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}\right)\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}-\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)}{\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}+\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)^3}$$
$$= \frac{s^2}{(1+sT_p)} y_p\left(1-\frac{y_p}{R}\right)\left(1-2\frac{y_p}{R}\right) \qquad (7.39)$$

$$\frac{\partial^2 y_p}{\partial T_p \partial K_p} = R\frac{\left(-\frac{sG_c e^{-s\tau_p}}{(1+sT_p)^2}\right)\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}\right)\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}-\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)}{\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}+\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)^3}$$

$$= -\frac{s}{K_p(1+sT_p)} y_p\left(1-\frac{y_p}{R}\right)\left(1-2\frac{y_p}{R}\right) \qquad (7.40)$$

Alternative Newton-Raphson gradient algorithms based on varying the assumptions used to determine the algorithm developed, may be defined. Two such algorithms are detailed in Section 7.3.2 and 7.3.3 (corresponding to the Gauss-Newton (2) algorithm and the Gauss-Newton (3) algorithm discussed in Section 7.2.2 and 7.2.3, respectively).

7.3.2 Theoretical development of the Newton-Raphson (2) algorithm

The Gauss-Newton (2) algorithm involved assuming that

$$\frac{\partial^2 J}{\partial \alpha_1 \partial \alpha_2} = \int e\frac{\partial^2 e}{\partial \alpha_1 \partial \alpha_2} + \frac{\partial e}{\partial \alpha_1}\frac{\partial e}{\partial \alpha_2} dt \approx \int \frac{\partial e}{\partial \alpha_1}\frac{\partial e}{\partial \alpha_2} dt \qquad (7.41)$$

in the development of equation (7.24). However, it is straightforward to calculate $\partial^2 e/\partial \alpha_1 \partial \alpha_2$ (by partially differentiating equations (7.26), (7.27) and (7.28)), facilitating the elimination of the assumption that this term may be neglected. The update values of the model parameters may be subsequently calculated, using a similar procedure to that given in Section 7.2.2, by solving the following relationship:

$$\begin{bmatrix} \int e\frac{\partial e}{\partial K_p}dt \\ \int e\frac{\partial e}{\partial T_p}dt \\ \int e\frac{\partial e}{\partial \tau_p}dt \end{bmatrix} = \begin{bmatrix} \int\left(\frac{\partial e}{\partial K_p}\right)^2 + e\frac{\partial^2 e}{\partial K_p^2}dt & \int\left(\frac{\partial e}{\partial T_p}\right)\left(\frac{\partial e}{\partial K_p}\right) + e\frac{\partial^2 e}{\partial T_p \partial K_p}dt & \int\left(\frac{\partial e}{\partial \tau_p}\right)\left(\frac{\partial e}{\partial K_p}\right) + e\frac{\partial^2 e}{\partial \tau_p \partial K_p}dt \\ \int\left(\frac{\partial e}{\partial T_p}\right)\left(\frac{\partial e}{\partial K_p}\right) + e\frac{\partial^2 e}{\partial T_p \partial K_p}dt & \int\left(\frac{\partial e}{\partial T_p}\right)^2 + e\frac{\partial^2 e}{\partial T_p^2}dt & \int\left(\frac{\partial e}{\partial T_p}\right)\left(\frac{\partial e}{\partial \tau_p}\right) + e\frac{\partial^2 e}{\partial \tau_p \partial T_p}dt \\ \int\left(\frac{\partial e}{\partial \tau_p}\right)\left(\frac{\partial e}{\partial K_p}\right) + e\frac{\partial^2 e}{\partial \tau_p \partial K_p}dt & \int\left(\frac{\partial e}{\partial T_p}\right)\left(\frac{\partial e}{\partial \tau_p}\right) + e\frac{\partial^2 e}{\partial T_p \partial \tau_p}dt & \int\left(\frac{\partial e}{\partial \tau_p}\right)^2 + e\frac{\partial^2 e}{\partial \tau_p^2}dt \end{bmatrix} \begin{bmatrix} \Delta K_m \\ \Delta T_m \\ \Delta \tau_m \end{bmatrix}$$

$$(7.42)$$

The following formulae have been developed for the evaluation of the required second partial derivatives in equation (7.42):

$$\frac{\partial^2 e}{\partial \tau_p^{\,2}} = R\frac{\left(\dfrac{s^2 G_c K_p e^{-s\tau_p}}{1+sT_p}\right)\left(1+\dfrac{G_c K_m}{1+sT_m}\right)\left(1+\dfrac{G_c K_m}{1+sT_m}-\dfrac{G_c K_m e^{-s\tau_m}}{1+sT_m}-\dfrac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)}{\left(1+\dfrac{G_c K_m}{1+sT_m}-\dfrac{G_c K_m e^{-s\tau_m}}{1+sT_m}+\dfrac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)^3}$$

$$= s^2 y_p\left(1-\frac{y_p}{R}+\frac{y_m}{R}\right)\left(1-2\frac{y_p}{R}\right) \qquad (7.43)$$

$$\frac{\partial^2 e}{\partial K_p^{\,2}} = R\frac{\left(-\dfrac{2G_c^{\,2} e^{-2s\tau_p}}{(1+sT_p)^2}\right)\left(1+\dfrac{G_c K_m}{1+sT_m}\right)}{\left(1+\dfrac{G_c K_m}{1+sT_m}-\dfrac{G_c K_m e^{-s\tau_m}}{1+sT_m}+\dfrac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)^3} = -\frac{2}{K_p^{\,2}} y_p \frac{y_p}{R}\left(1-\frac{y_p}{R}+\frac{y_m}{R}\right) \qquad (7.44)$$

$$\frac{\partial^2 e}{\partial T_p^{\,2}} = R\frac{\left(\dfrac{2s^2 G_c K_p e^{-s\tau_p}}{(1+sT_p)^3}\right)\left(1+\dfrac{G_c K_m}{1+sT_m}\right)\left(1+\dfrac{G_c K_m}{1+sT_m}-\dfrac{G_c K_m e^{-s\tau_m}}{1+sT_m}\right)}{\left(1+\dfrac{G_c K_m}{1+sT_m}-\dfrac{G_c K_m e^{-s\tau_m}}{1+sT_m}+\dfrac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)^3}$$

$$= \frac{2s^2}{(1+sT_p)^2} y_p\left(1-\frac{y_p}{R}+\frac{y_m}{R}\right)\left(1-\frac{y_p}{R}\right) \qquad (7.45)$$

$$\frac{\partial^2 e}{\partial K_p \partial \tau_p} = R\frac{\left(-\dfrac{s G_c e^{-s\tau_p}}{1+sT_p}\right)\left(1+\dfrac{G_c K_m}{1+sT_m}\right)\left(1+\dfrac{G_c K_m}{1+sT_m}-\dfrac{G_c K_m e^{-s\tau_m}}{1+sT_m}-\dfrac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)}{\left(1+\dfrac{G_c K_m}{1+sT_m}-\dfrac{G_c K_m e^{-s\tau_m}}{1+sT_m}+\dfrac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)^3}$$

$$= -\frac{s}{K_p} y_p\left(1-\frac{y_p}{R}+\frac{y_m}{R}\right)\left(1-2\frac{y_p}{R}\right) \qquad (7.46)$$

$$\frac{\partial^2 e}{\partial T_p \partial \tau_p} = R\frac{\left(\frac{s^2 G_c K_p e^{-s\tau_p}}{(1+sT_p)^2}\right)\left(1+\frac{G_c K_m}{1+sT_m}\right)\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}-\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)}{\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}+\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)^3}$$

$$= \frac{s^2}{(1+sT_p)} y_p \left(1-\frac{y_p}{R}+\frac{y_m}{R}\right)\left(1-2\frac{y_p}{R}\right) \tag{7.47}$$

$$\frac{\partial^2 e}{\partial T_p \partial K_p} = R\frac{\left(-\frac{s G_c e^{-s\tau_p}}{(1+sT_p)^2}\right)\left(1+\frac{G_c K_m}{1+sT_m}\right)\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}-\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)}{\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}+\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)^3}$$

$$= -\frac{s}{K_p(1+sT_p)} y_p \left(1-\frac{y_p}{R}+\frac{y_m}{R}\right)\left(1-2\frac{y_p}{R}\right) \tag{7.48}$$

7.3.3 Theoretical development of the Newton-Raphson (3) algorithm

The Gauss-Newton (3) algorithm involved assuming that

$$\frac{\partial^2 J}{\partial \alpha_1 \partial \alpha_2} \approx \int e\frac{\partial^2 e_1}{\partial \alpha_1 \partial \alpha_2} + \frac{\partial e_1}{\partial \alpha_1}\frac{\partial e_1}{\partial \alpha_2} dt \approx \int \frac{\partial e_1}{\partial \alpha_1}\frac{\partial e_1}{\partial \alpha_2} \tag{7.49}$$

with $e_1 = 0.5(y_p + y_m)$, to develop equation (7.30). However, it is straightforward to calculate $\partial^2 e_1/\partial\alpha_1\,\partial\alpha_2$ (by partially differentiating equations (7.31), (7.32) and (7.33)), facilitating the elimination of the assumption that this term may be neglected. The update values of the model parameters may be subsequently calculated, using a similar procedure to that given in Section 7.2.3, by solving the following relationship:

$$\begin{bmatrix} \int e\frac{\partial e_1}{\partial K_p}dt \\ \int e\frac{\partial e_1}{\partial T_p}dt \\ \int e\frac{\partial e_1}{\partial \tau_p}dt \end{bmatrix} = \begin{bmatrix} \int\left(\frac{\partial e_1}{\partial K_p}\right)^2 + e\frac{\partial^2 e_1}{\partial K_p^2}dt & \int\left(\frac{\partial e_1}{\partial T_p}\right)\left(\frac{\partial e_1}{\partial K_p}\right) + e\frac{\partial^2 e_1}{\partial T_p \partial K_p}dt & \int\left(\frac{\partial e_1}{\partial \tau_p}\right)\left(\frac{\partial e_1}{\partial K_p}\right) + e\frac{\partial^2 e_1}{\partial \tau_p \partial K_p}dt \\ \int\left(\frac{\partial e_1}{\partial T_p}\right)\left(\frac{\partial e_1}{\partial K_p}\right) + e\frac{\partial^2 e_1}{\partial T_p \partial K_p}dt & \int\left(\frac{\partial e_1}{\partial T_p}\right)^2 + e\frac{\partial^2 e_1}{\partial T_p^2}dt & \int\left(\frac{\partial e_1}{\partial T_p}\right)\left(\frac{\partial e_1}{\partial \tau_p}\right) + e\frac{\partial^2 e_1}{\partial \tau_p \partial T_p}dt \\ \int\left(\frac{\partial e_1}{\partial \tau_p}\right)\left(\frac{\partial e_1}{\partial K_p}\right) + e\frac{\partial^2 e_1}{\partial \tau_p \partial K_p}dt & \int\left(\frac{\partial e_1}{\partial T_p}\right)\left(\frac{\partial e_1}{\partial \tau_p}\right) + e\frac{\partial^2 e_1}{\partial T_p \partial \tau_p}dt & \int\left(\frac{\partial e_1}{\partial \tau_p}\right)^2 + e\frac{\partial^2 e_1}{\partial \tau_p^2}dt \end{bmatrix} \begin{bmatrix} K_e \\ T_e \\ \tau_e \end{bmatrix}$$

(7.50)

The following formulae have been developed for the evaluation of some of the required second partial derivatives in equation (7.50):

$$\frac{\partial^2 e_1}{\partial \tau_p^{\,2}} = R\frac{\left(\frac{s^2 G_c K_p e^{-s\tau_p}}{1+sT_p}\right)\left(1+\frac{G_c K_m}{1+sT_m}-2\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}\right)\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}-\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)}{\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}+\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)^3}$$

$$= s^2 y_p\left(1-\frac{y_p}{R}-\frac{y_m}{R}\right)\left(1-2\frac{y_p}{R}\right) \qquad (7.51)$$

$$\frac{\partial^2 e_1}{\partial K_p^{\,2}} = R\frac{\left(-\frac{2G_c^{\,2} e^{-2s\tau_p}}{(1+sT_p)^2}\right)\left(1+\frac{G_c K_m}{1+sT_m}-2\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}\right)}{\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}+\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)^3} = -\frac{2}{K_p^{\,2}} y_p \frac{y_p}{R}\left(1-\frac{y_p}{R}-\frac{y_m}{R}\right)$$

$$(7.52)$$

$$\frac{\partial^2 e_1}{\partial T_p^{\,2}} = R\frac{\left(\frac{2s^2 G_c K_p e^{-s\tau_p}}{(1+sT_p)^3}\right)\left(1+\frac{G_c K_m}{1+sT_m}-2\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}\right)}{\left(1+\frac{G_c K_m}{1+sT_m}-\frac{G_c K_m e^{-s\tau_m}}{1+sT_m}+\frac{G_c K_p e^{-s\tau_p}}{1+sT_p}\right)^3}$$

$$= \frac{2s^2}{(1+sT_p)^2} y_p\left(1-\frac{y_p}{R}-\frac{y_m}{R}\right)\left(1-\frac{y_p}{R}\right) \qquad (7.53)$$

Similar terms may be determined for $\partial^2 e_1/\partial\alpha_1\,\partial\alpha_2$, $\alpha_1 \neq \alpha_2$.

7.3.4 Algorithm representations

The block diagram representation of the algorithms for updating the model time delay only is provided in Figure 7.9. The block diagram of the three algorithms used differs only in the sensitivity functions; $v = y_p$, e and e_1 for the Newton-Raphson (1), Newton-Raphson (2) and Newton-Raphson (3) algorithms, respectively.

The block diagram representation of the algorithms for updating the model gain or model time constant only, assuming a FOLPD model, is provided in Figure 7.10. In the sensitivity function in the diagram, $\omega = K_p$ when the model gain is to be updated and $\omega = T_p$ when the model time constant is to be updated. The code for v is identical to that in Figure 7.9.

The block diagram representation of the sensitivity functions for the Newton-Raphson (1) algorithm is provided in Figure 7.11; $K_m{}', T_m{}'$ and $\tau_m{}'$ are approximations for the (unknown) process parameter values used in the sensitivity functions (the explicit dependence of the sensitivity functions on the process parameter values is shown in equations (7.51) to (7.53), for example). The block diagram representation of the sensitivity functions for the other two Newton-Raphson algorithms may be specified in a similar manner (O'Dwyer (1996i)). The block diagram representation of the algorithms to facilitate the updating of all the model parameters simultaneously may be subsequently defined, as may the representation of the algorithms that allow updating of the parameters of an arbitrary order model plus time delay.

It is interesting that the calculation of the $\partial^2 y_p / \partial \tau_p{}^2$ sensitivity function (Figure 7.11) involves the use of a squared derivative term; the effect of noise on the operation of this calculation will therefore be greater than its effect on the corresponding sensitivity function ($\partial y_p / \partial \tau_p$) for the Gauss-Newton (1) algorithm (which is, of course, quite sensitive itself to the presence of noise).

Figure 7.9: Representation of the Newton-Raphson algorithms, for time delay estimation.

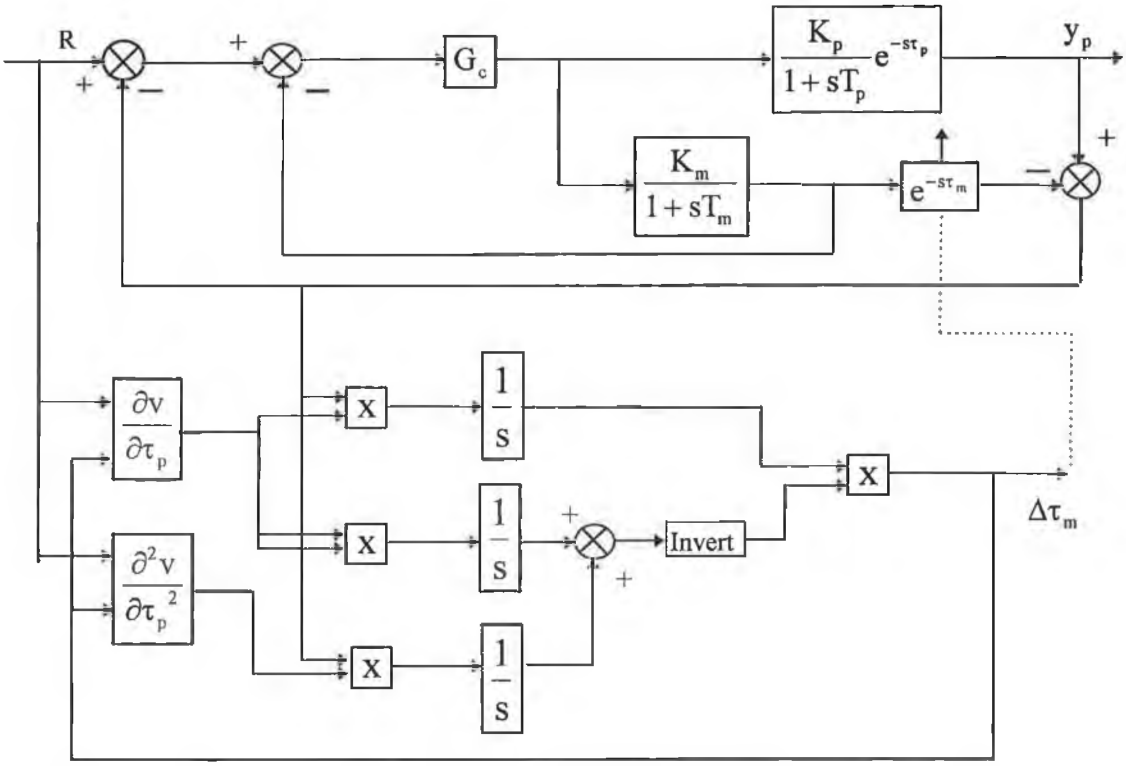


Figure 7.10: Representation of the Newton-Raphson algorithms, for model gain or model time constant estimation.

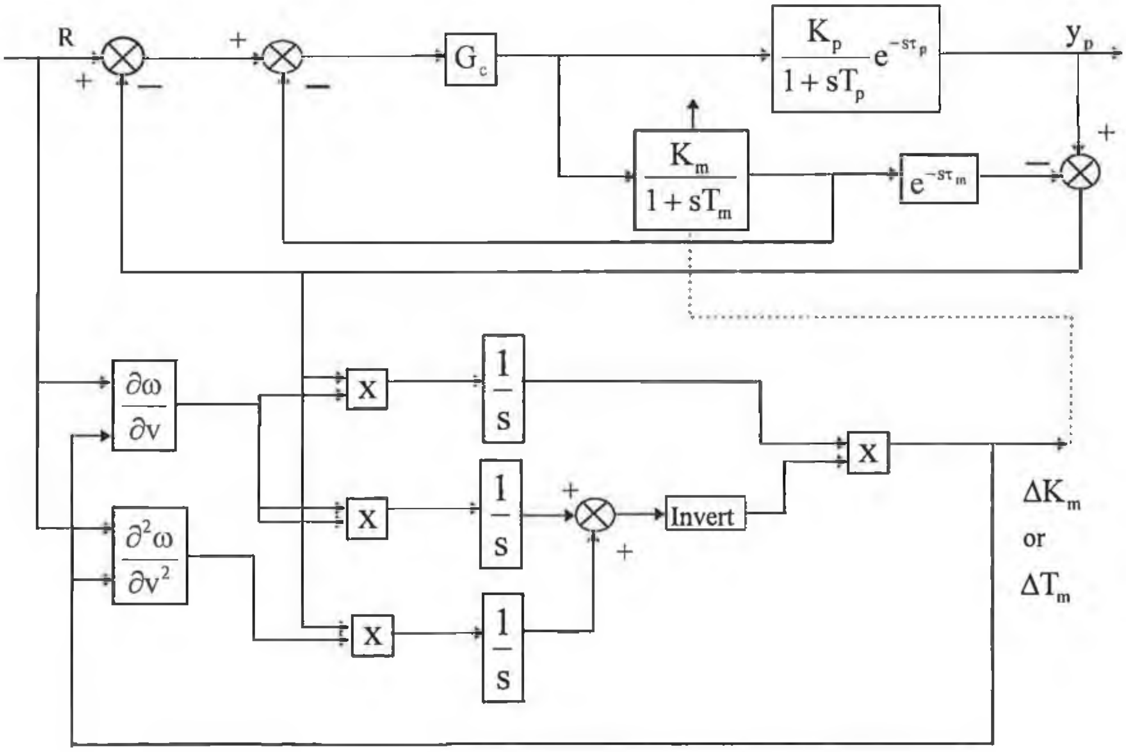
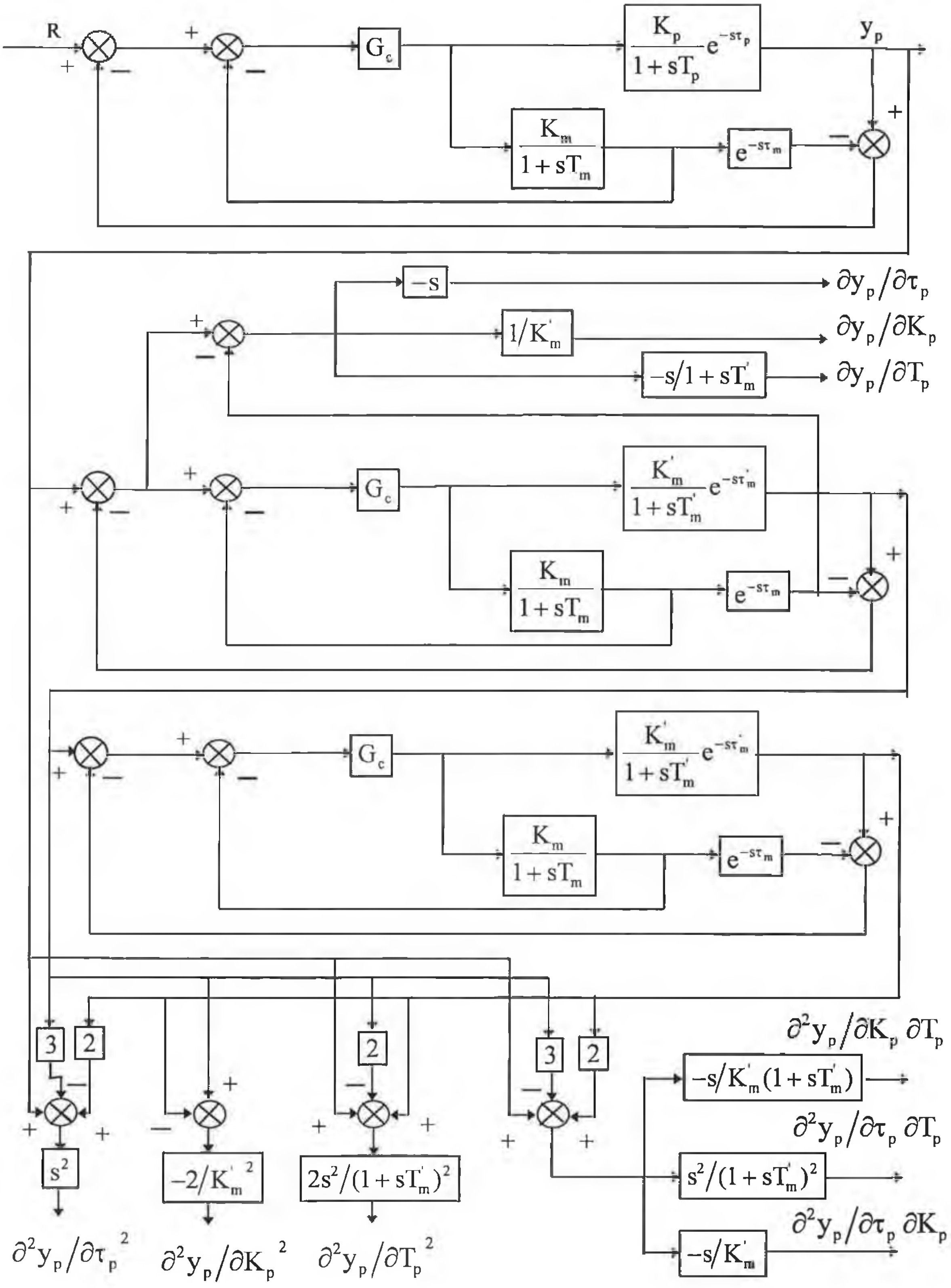


Figure 7.11: Representation of the sensitivity functions for the Newton-Raphson (1) algorithm.



7.4 Parameter estimation - simulation results

The algorithms defined have been simulated, for updating all of the parameters separately, using the SIMULINK package. It has been decided to simulate the algorithms for seven process/model combinations, in a Smith predictor structure. The processes considered are the same as those identified in Chapter 4; their transfer functions are provided in equations (4.57) to (4.63), inclusive. They include high order, underdamped and non-minimum phase processes, which were modelled by equivalently ordered models or mismatched FOLPD and SOSPD models, as appropriate. The FOLPD and SOSPD models were obtained from the two stage frequency domain identification technique of Chapter 4. The PI and PID primary controllers used are specified to be robust to the possible process/model parameter mismatches considered; the design procedures for the controllers include the robust synthesis procedure of Morari and Zafiriou (1989). In all cases, the maximum variation of the process parameters must be known *a priori*.

In each simulation, the excitation signal at the servo input is of band limited white noise form; such a signal was determined to be sufficiently exciting so that appropriate parameter updating is achieved. In all cases, the individual model parameters are updated at discrete intervals using a dedicated s-function in SIMULINK; the gradient algorithm implementations, which are in continuous time (as shown in Figures 7.3 to 7.11) are also effectively set up in continuous time in the SIMULINK environment (by choosing a small step size for the simulations). The approximations for the process parameters in the sensitivity functions (labelled K_m', T_m' and τ_m' in Figures 7.5 and 7.11) are updated at the same rate as the model parameters; the realisable sensitivity functions produced are the best approximations to the 'ideal' sensitivity functions.

7.4.1 Time delay estimation

The six algorithms for updating the time delay have been simulated individually, for each of the process-model combinations outlined above. Representative simulation results are provided in Cases 1 to 8, in which all of the gradient algorithms defined are used at some stage; further supplementary simulation results are provided by O'Dwyer (1996j).

Case 1: $G_m e^{-s\tau_m} = 2e^{-1.4s}/(1+0.7s)$, $G_c = 1.75(1+1/0.7s)$. In Figures 7.12 and 7.13, τ_p = 1.2 seconds and $G_p = G_m$; in Figures 7.14 and 7.15, τ_p = 1.6 seconds and $G_p = G_m$.

Figure 7.12: Time delay updating

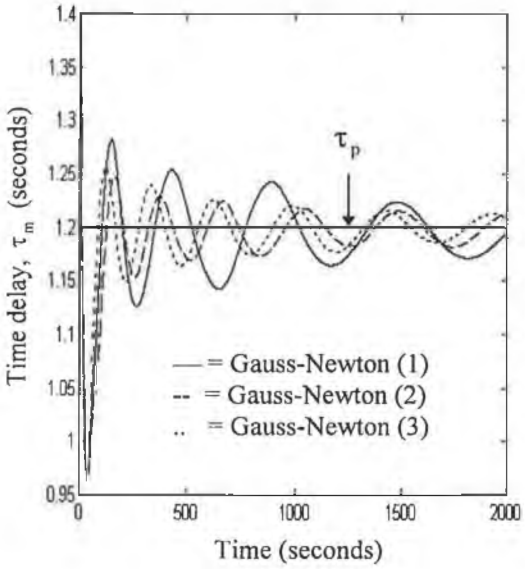


Figure 7.13: Time delay updating

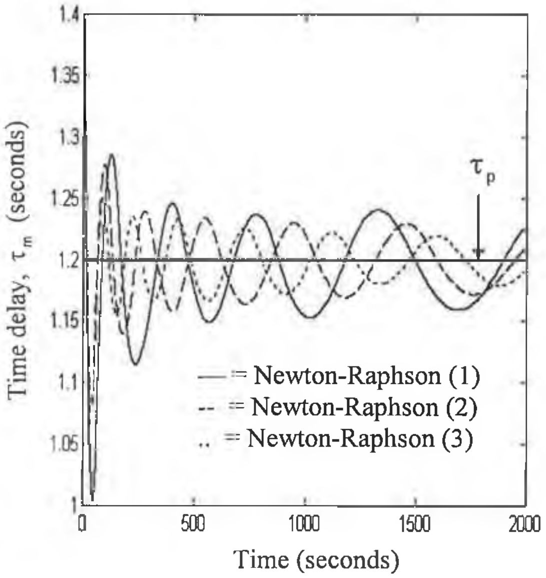


Figure 7.14 : Time delay updating

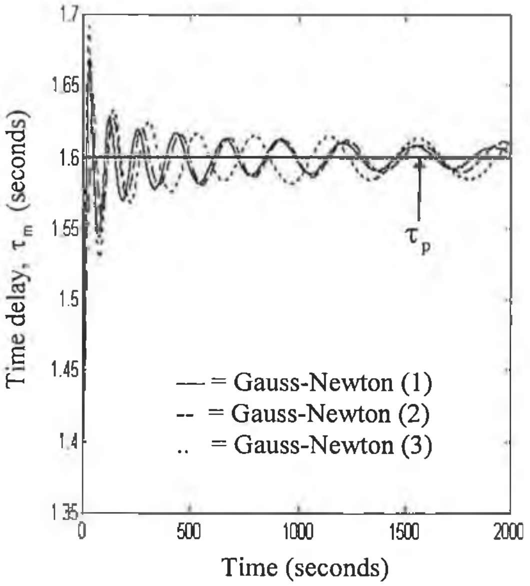
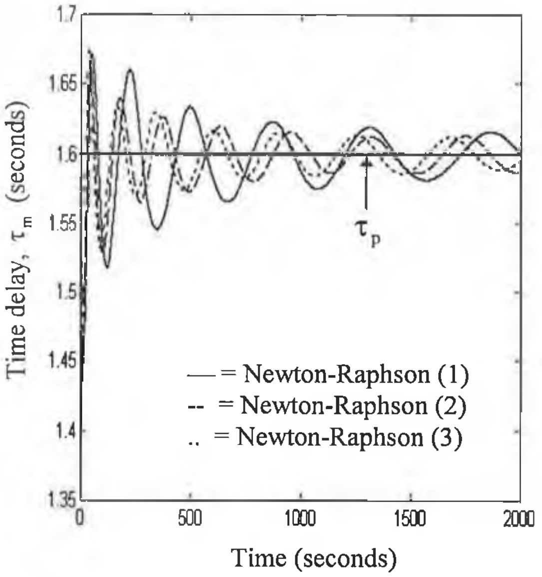


Figure 7.15: Time delay updating



Figures 7.12 to 7.15 show that the algorithms facilitate a reduction in mismatch between the process time delay and the model time delay.

Case 2: $G_m e^{-s\tau_m} = 2e^{-s}/(1+4.5+4.5s^2), G_c = 1.17(1+1/4.07s+2.73s/(1+0.5s))$. In Figure 7.16, $\tau_p = 0.7$ seconds and $G_p = G_m$; in Figure 7.17, $\tau_p = 1.3$ seconds and $G_p = G_m$.

Figure 7.16: Time delay updating

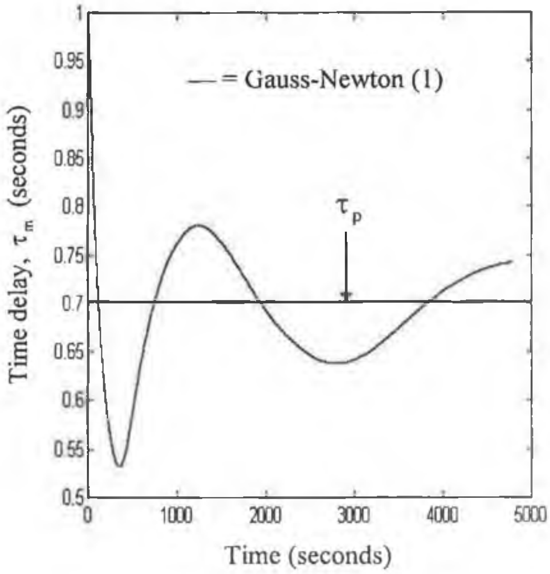
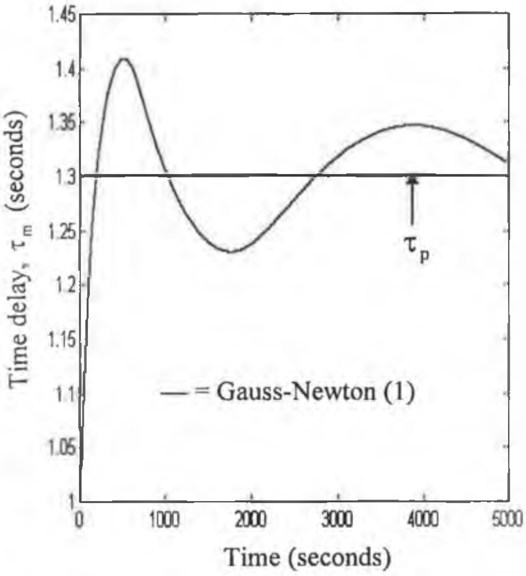


Figure 7.17: Time delay updating



Case 3: $G_m e^{-s\tau_m} = 2e^{-s}/(1+18s+137s^2+567s^3+1403s^4+2103s^5+1846s^6+856s^7+158s^8)$, $G_c = 2.14(1+1/9.75s+3.31s/(1+0.61s))$. In Figure 7.18, $\tau_p = 0.7$ seconds and $G_p = G_m$; in Figure 7.19, $\tau_p = 1.3$ seconds and $G_p = G_m$.

Figure 7.18: Time delay updating

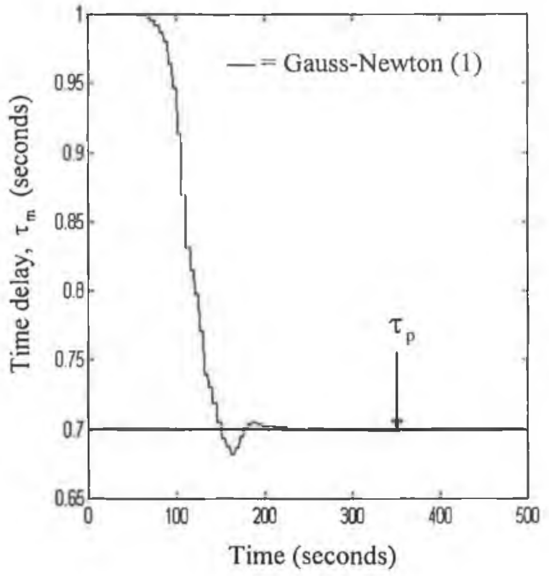
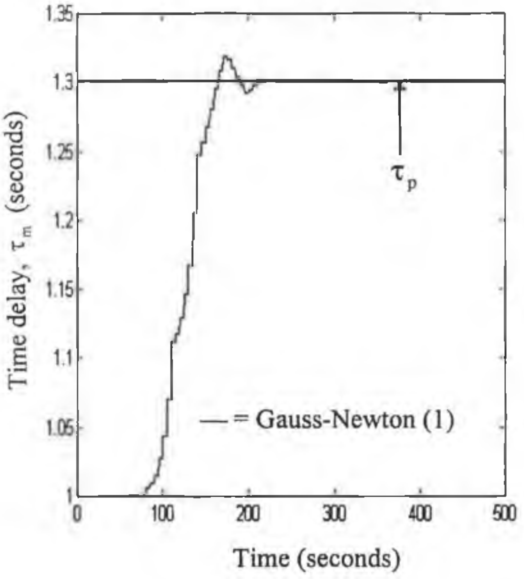


Figure 7.19: Time delay updating



A similar comment to that made in Case 1 applies to the simulation results in Figures 7.16 to 7.19.

Case 4: $G_m e^{-s\tau_m} = 2e^{-1.4s}/(1+0.7s)$, $G_c = 1.75(1+1/0.7s)$. In Figures 7.20 and 7.21, τ_p = 1.2 seconds and $G_p = 1.6/(1+0.5s)$; in Figures 7.22 and 7.23, τ_p = 1.6 seconds and $G_p = 2.4/(1+0.9s)$.

Figure 7.20: Time delay updating

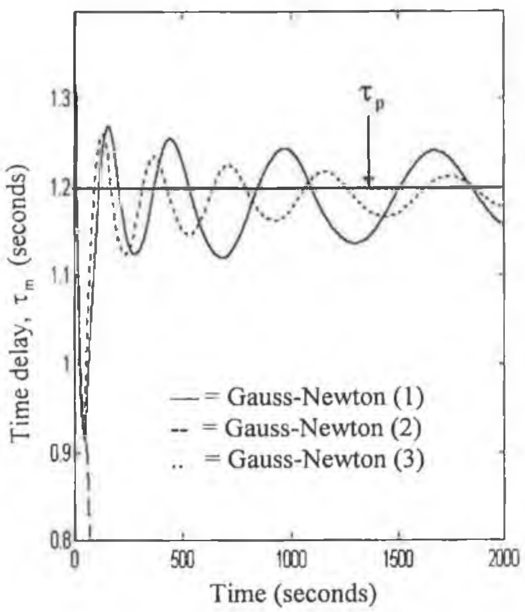


Figure 7.21: Time delay updating

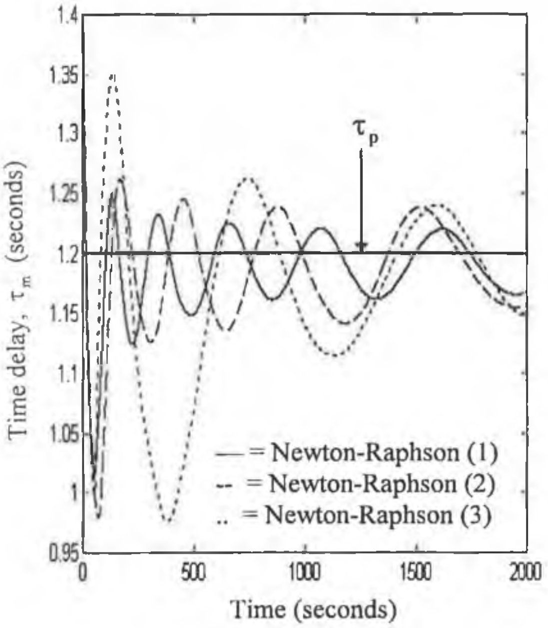


Figure 7.22: Time delay updating

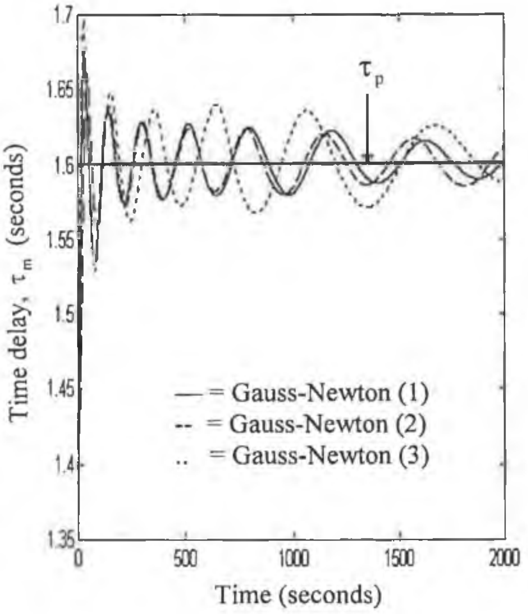
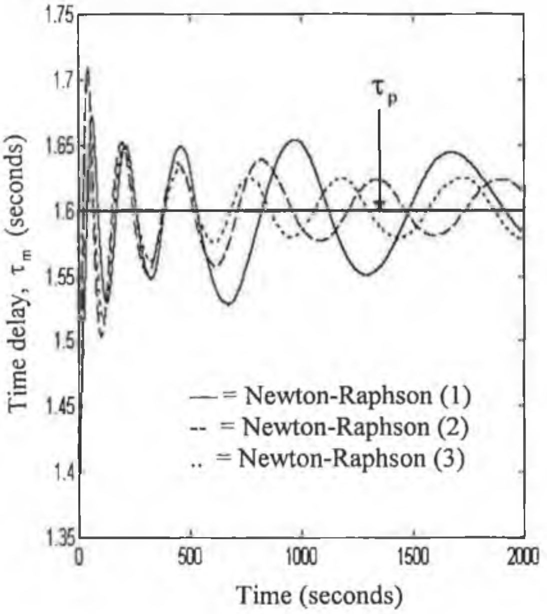


Figure 7.23: Time delay updating



Figures 7.20 to 7.23 show that most of the algorithms facilitate a reduction in mismatch between the process and the model. These are significant results, as the non-delay process and model parameters are different.

Case 5: $G_m e^{-s\tau_m} = 2e^{-s}/(1+4.5s+4.5s^2)$, $G_c = 1.17(1+1/4.07s+2.73s/(1+0.5s))$. In Figures 7.24 and 7.26, $\tau_p = 0.7$ seconds and $G_p = 1.2/(1+3.1s+3.1s^2)$; in Figures 7.25 and 7.27, $\tau_p = 1.3$ seconds and $G_p = 2.8/(1+6.1s+6.1s^2)$.

Figure 7.24: Time delay updating

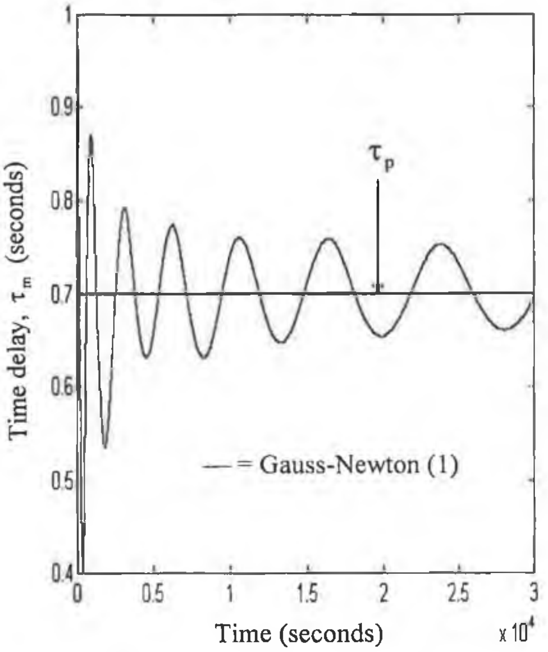


Figure 7.25: Time delay updating

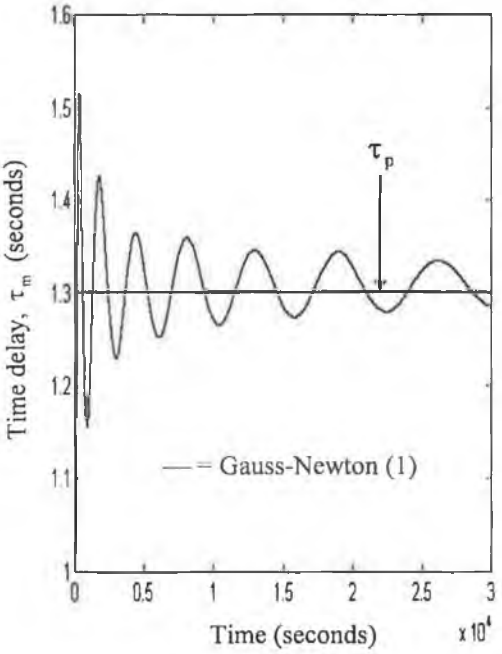


Figure 7.26: Time delay updating

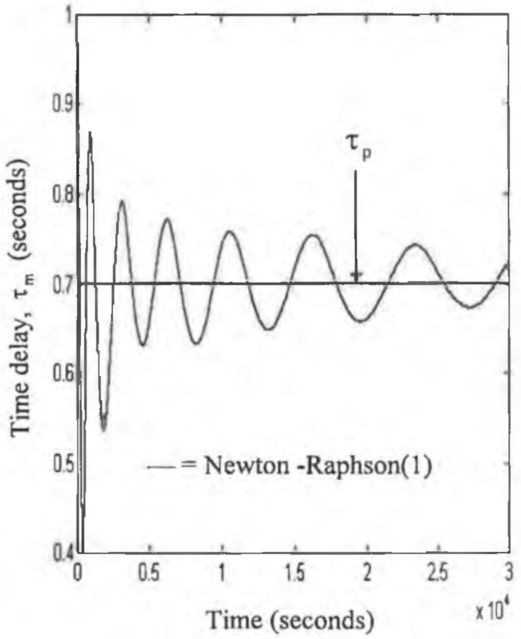
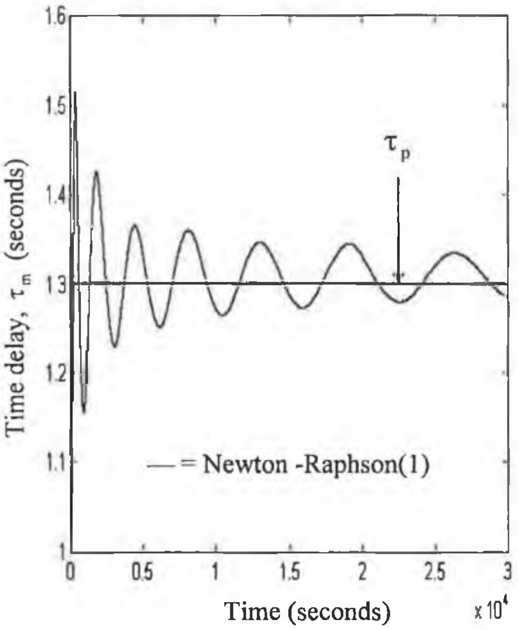


Figure 7.27: Time delay updating



A similar comment to that made in Case 4 applies to the simulation results in Figures 7.24 to 7.27 above.

Case 6: $G_m e^{-s\tau_m} = 1.96e^{-1.73s}/(1+4.65s)$; the process corresponding to this model is $2e^{-s}/(1+4.5s+4.5s^2)$, with $G_c = 1.19(1+1/4.65s)$. In Figures 7.28, 7.30 and 7.32, τ_p = 0.7 seconds and $G_p = 1.2/(1+3.1s+3.1s^2)$; in Figures 7.29, 7.31 and 7.33, τ_p = 1.3 seconds and $G_p = 2.8/(1+6.1s+6.1s^2)$.

Figure 7.28: Time delay updating

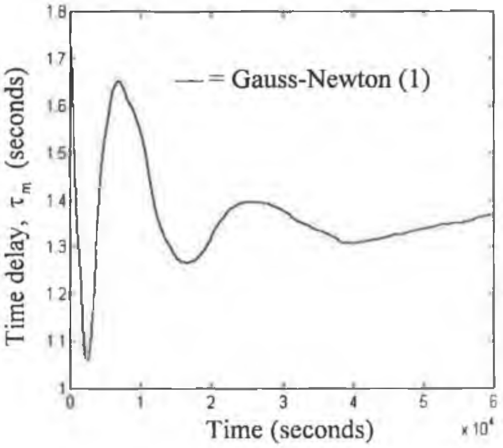


Figure 7.29: Time delay updating

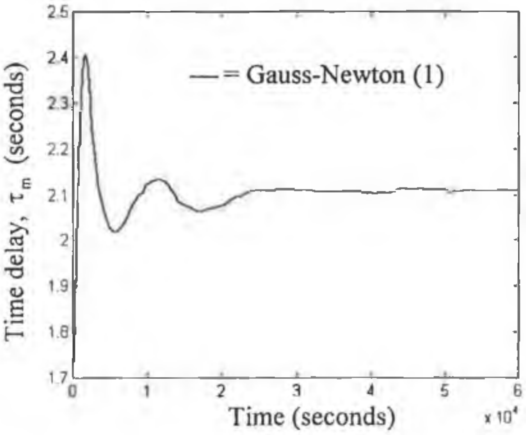


Figure 7.30: Time delay updating

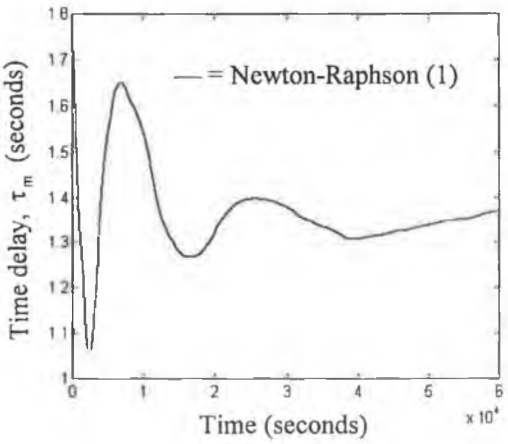


Figure 7.31: Time delay updating

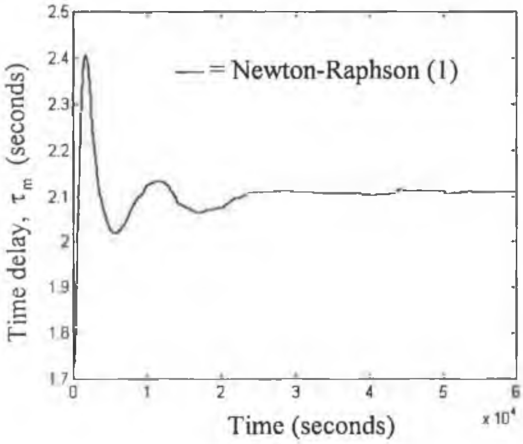


Figure 7.32: Time delay updating

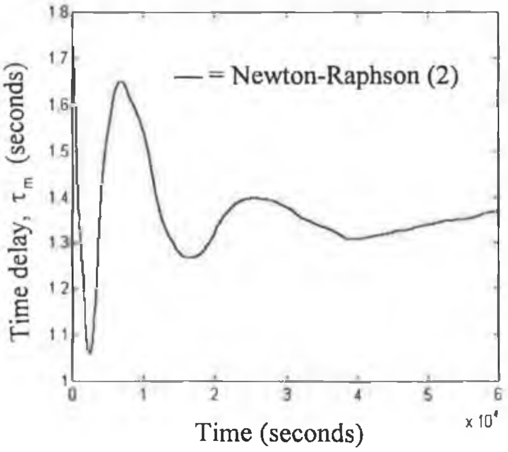
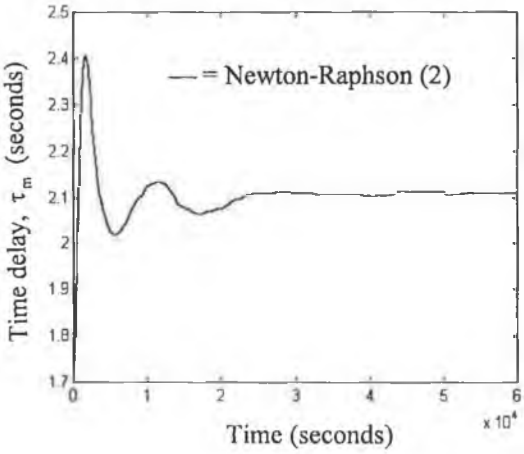
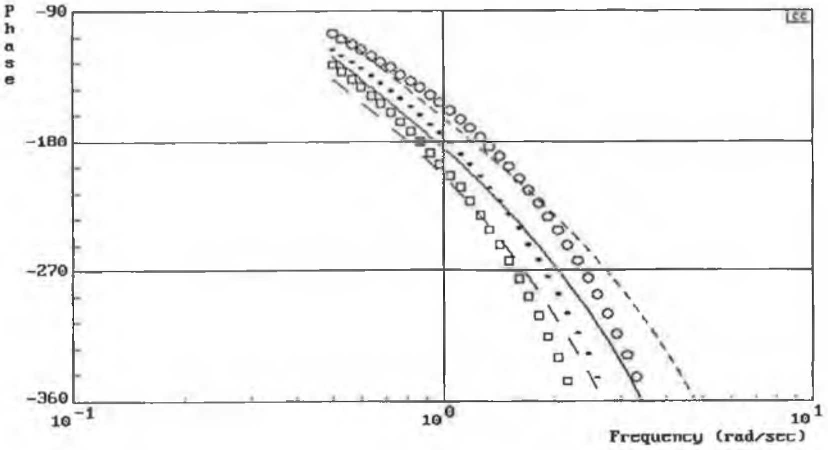


Figure 7.33: Time delay updating



Figures 7.28 to 7.33 show that when $\tau_p = 0.7$ seconds, τ_m approximately converges to 1.35 seconds and when $\tau_p = 1.3$ seconds, τ_m approximately converges to 2.11 seconds. It is difficult to sensibly compare the goodness of fit between the process and the model in the time or frequency domains, as all the parameters have been changed in the process and only the time delay is updated in the model; a polar plot shows poor fitting of the processes and their corresponding models for this reason. However, if phase plots of the processes and models are obtained at higher frequencies (Figure 7.34), reasonable fitting between the process and model is seen; this result implies that the model time delay estimates are appropriate, if it is desired to reduce the mismatch between the process and the model, as the time delay will be the dominant influence on the phase plot at higher frequencies. However, it is normally desirable when using a Smith predictor to reduce the mismatch between the process and model time delays; the matching of the process and the model at higher frequencies means that the difference between the process and the model, fed back in the Smith predictor, is small at these frequencies. This is not desirable, bearing in mind the large mismatch between the process and model time delays. Thus, the gradient algorithms may not be suited for updating the time delay in a Smith predictor structure, if the process and model orders are different.

Figure 7.34: Phase plot of processes and their models



with - .. $G_p e^{-s\tau_p} = \frac{2e^{-s}}{1+4.5s+4.5s^2}$, • .. $G_m e^{-s\tau_m} = \frac{1.96e^{-1.73s}}{1+4.65s}$, -- .. $G_p e^{-s\tau_p} = \frac{1.2e^{-0.7s}}{1+3.1s+3.1s^2}$,

o .. $G_m e^{-s\tau_m} = \frac{1.96e^{-1.35s}}{1+4.65s}$, —— .. $G_p e^{-s\tau_p} = \frac{2.8e^{-1.3s}}{1+6.1s+6.1s^2}$ and □ .. $G_m e^{-s\tau_m} = \frac{1.96e^{-2.11s}}{1+4.65s}$.

Case 7: $G_m e^{-s\tau_m} = 2.01e^{-0.99s}/(1+6.32s+8.25s^2)$. The process corresponding to this model is $(2+4.5s)e^{-s}/(1+8.5s+22.5s^2+18s^3)$. $G_c = 3.64(1+1/6.11s+1.14s/(1+0.21s))$. In Figures 7.35, 7.37 and 7.39, $G_p = (1.2+3.1s)/(1+5.9s+15.7s^2+12.6s^3)$ and τ_p = 0.7 seconds; in Figures 7.36, 7.38 and 7.40, $G_p = (2.8+6.1s)/(1+11s+29.3s^2+23.4s^3)$ and τ_p = 1.3 seconds.

Figure 7.35: Time delay updating

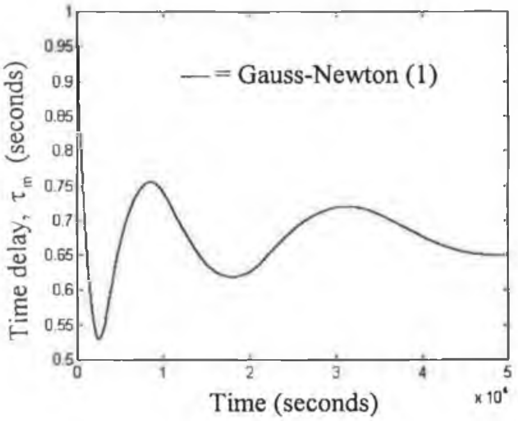


Figure 7.36: Time delay updating

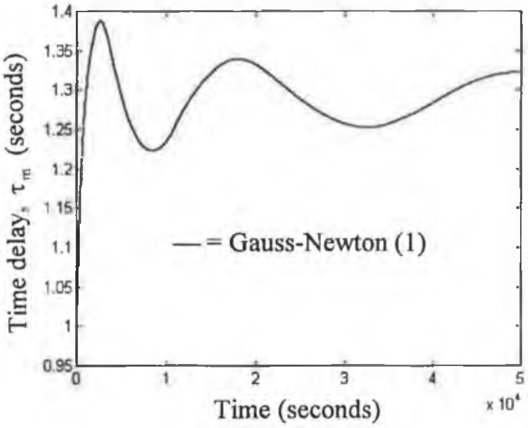


Figure 7.37: Time delay updating

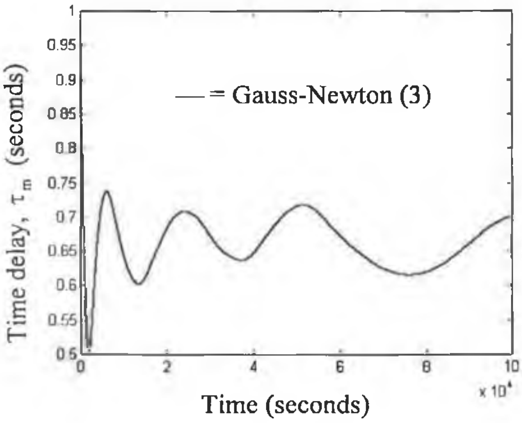


Figure 7.38: Time delay updating

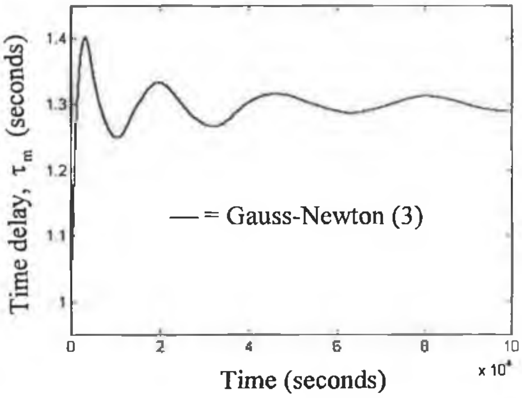


Figure 7.39: Time delay updating

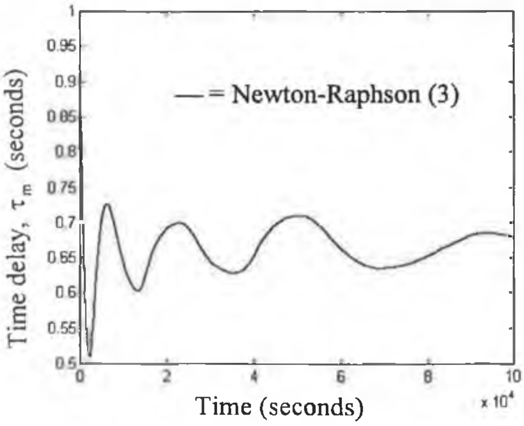
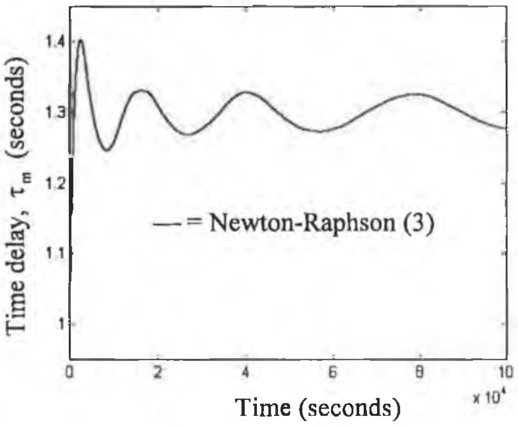
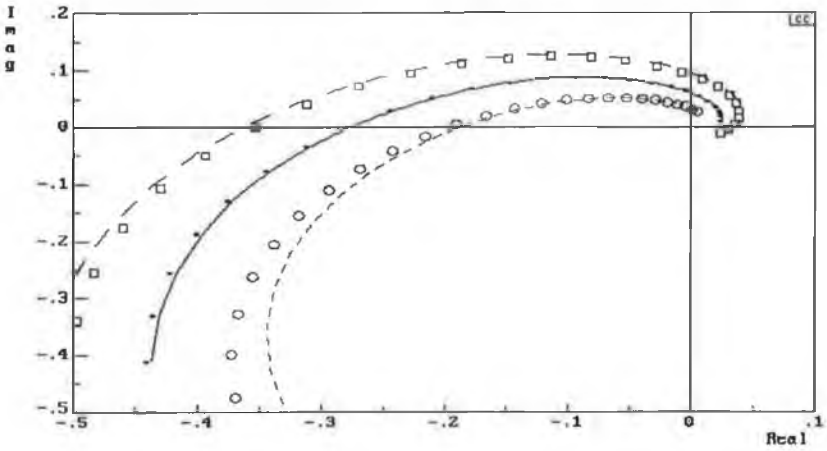


Figure 7.40: Time delay updating



Figures 7.35 to 7.40 show that when $\tau_p = 0.7$ seconds, τ_m approximately converges to 0.67 seconds and when $\tau_p = 1.3$ seconds, τ_m approximately converges to 1.29 seconds. Good fitting between the processes and models are seen, particularly at higher frequencies, when polar plots of the processes and models are obtained (Figure 7.41).

Figure 7.41: Polar plots of processes and their models



- .. $G_p e^{-s\tau_p} = \dfrac{(2+4.5s)e^{-s}}{1+8.5s+22.5s^2+18s^3}$, ● .. $G_m e^{-s\tau_m} = \dfrac{2.01e^{-0.99s}}{1+6.32s+8.25s^2}$, -- .. $G_p e^{-s\tau_p} = \dfrac{(1.2+3.1s)e^{-0.7s}}{1+5.9s+15.7s^2+12.6s^3}$

O .. $G_m e^{-s\tau_m} = \dfrac{2.01e^{-0.67s}}{1+6.32s+8.25s^2}$, —— .. $G_p e^{-s\tau_p} = \dfrac{(2.8+6.1s)e^{-1.3s}}{1+11s+29.3s^2+23.4s^3}$, □.. $G_m e^{-s\tau_m} = \dfrac{2.01e^{-1.29s}}{1+6.32s+8.25s^2}$

Case 8: $G_m e^{-s\tau_m} = 1.96e^{-1.84s}/(1+6.7s)$; the process corresponding to this model is $(2+4.5s)e^{-s}/(1+8.5s+22.5s^2+18s^3)$ with $G_c = 6.84(1+1/6.7s)$. In Figures 7.42, 7.44 and 7.46, $G_p = (1.2+3.1s)/(1+5.9s+15.7s^2+12.6s^3)$ and $\tau_p = 0.7$ seconds; in Figures 7.43, 7.45 and 7.47, $G_p = (2.8+6.1s)/(1+11s+29.3s^2+23.4s^3)$ and $\tau_p = 1.3$ seconds.

Figure 7.42: Time delay updating

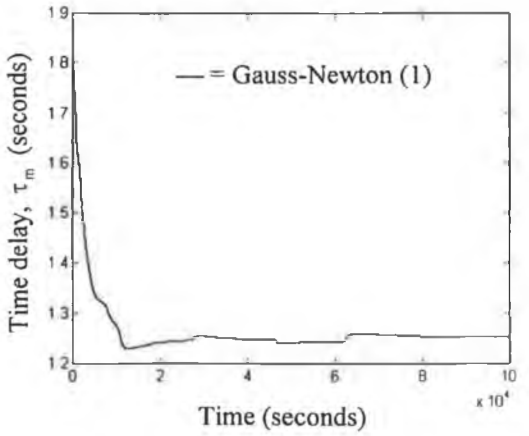


Figure 7.43: Time delay updating

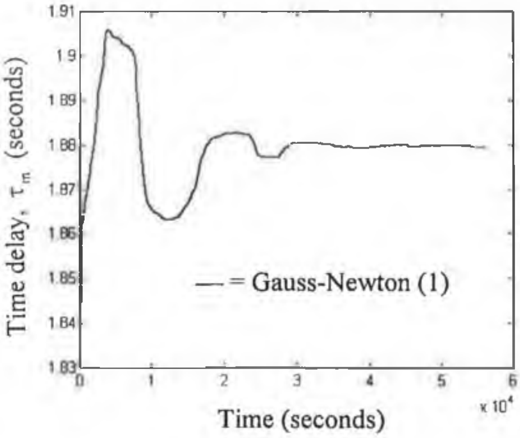


Figure 7.44: Time delay updating

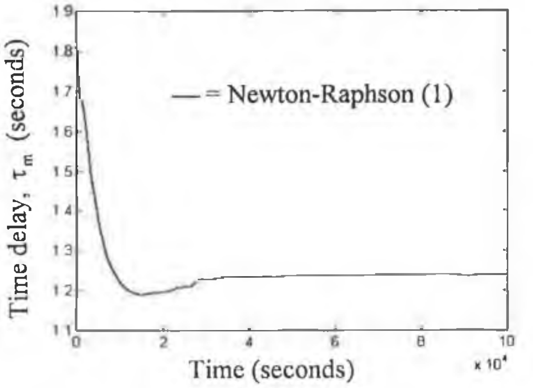


Figure 7.45: Time delay updating

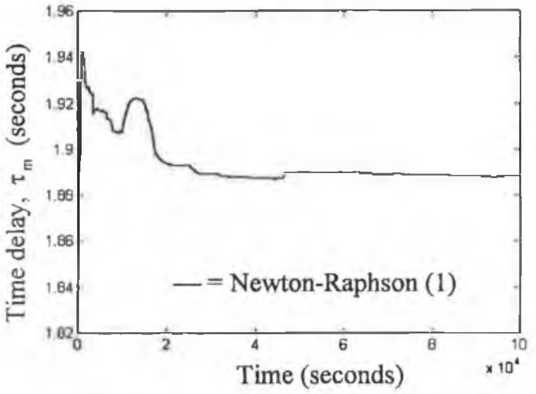


Figure 7.46: Time delay updating

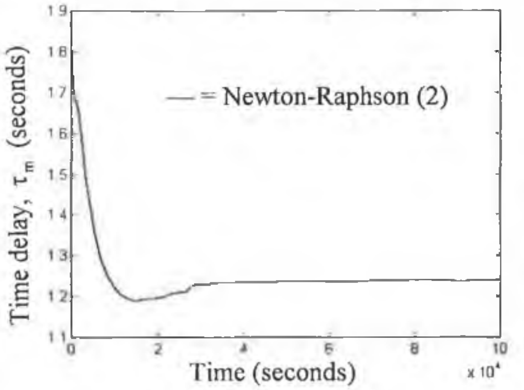
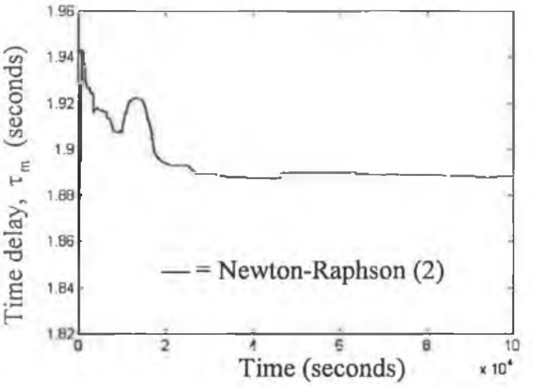
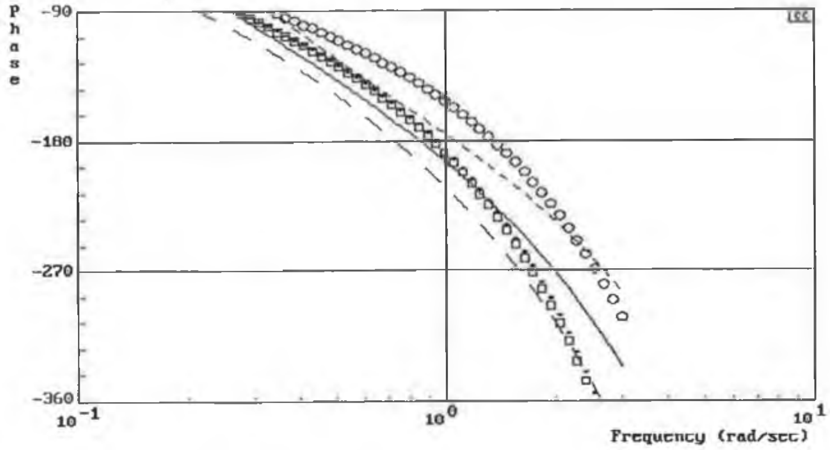


Figure 7.47: Time delay updating



Figures 7.42 to 7.47 show that when $\tau_p = 0.7$ seconds, τ_m approximately converges to 1.25 seconds and when $\tau_p = 1.3$ seconds, τ_m approximately converges to 1.88 seconds. As in Case 6, good fitting between the processes and models is seen if phase plots of the processes and models are obtained at higher frequencies (Figure 7.48).

Figure 7.48: Phase plots of processes and their models



- .. $G_pe^{-s\tau_p} = \frac{(2+4.5s)e^{-s}}{1+8.5s+22.5s^2+18s^3}$, ● .. $G_me^{-s\tau_m} = 1.96e^{-1.84s}/(1+6.7s)$, -- .. $G_pe^{-s\tau_p} = \frac{(1.2+3.1s)e^{-0.7s}}{1+5.9s+15.7s^2+12.6s^3}$

O .. $G_me^{-s\tau_m} = 1.96e^{-1.25s}/(1+6.7s)$, — — .. $G_pe^{-s\tau_p} = \frac{(2.8+6.1s)e^{-1.3s}}{1+11s+29.3s^2+23.4s^3}$, ☐ .. $G_me^{-s\tau_m} = 1.96e^{-1.89s}/(1+6.7s)$

Overall, the full panorama of simulation results show that when the order of the process equals that of the model, the mismatch between the model time delay and the process time delay is significantly reduced, irrespective of the match between the process and model parameters. When the order of the model differs from that of the process, then the model delay is updated to a final value in the simulations taken.

The performance of the six algorithms is more difficult to compare, though it is obvious that there is little to be gained (at least in the simulations taken) in using a Newton-Raphson algorithm instead of a Gauss-Newton algorithm. It is evident that, on balance, and taking the full panorama of simulation results obtained (including the supplementary results provided by O'Dwyer (1996j)), the Gauss-Newton (1) time delay updating algorithm is the most appropriate algorithm to use, with the Gauss-Newton (2) algorithm being the least appropriate one to use. This conclusion has been reached based on the speed and reliability of the convergence of the relevant parameters over the full range of simulations taken. It may be shown analytically by checking the validity of the assumptions used in the algorithm for the application, that a Newton-Raphson algorithm had advantages over a Gauss-Newton algorithm as the process-model mismatch increases. Unfortunately, because of the complexity of the equations developed to check the validity of the assumptions, and their dependence on knowledge of the (unknown) process parameters, it was not possible to arrive at a general conclusion as to the algorithms' relative performance, as indeed the simulation results discussed above reveal.

It is interesting that it takes a long time for the model time delay to converge to the process time delay in most cases, even when the order of the process and model are the same. The oscillatory convergence pattern is a factor in this disappointingly slow convergence rate; an alteration in the learning rate of the gradient algorithms (discussed in Chapter 2, Section 2.2.4) would improve this situation.

7.4.2 Estimation of the non-delay parameters

The six algorithms for separately updating the gain and the time constant have been simulated individually, for the process-model combination in which both the process and model are in FOLPD form. It was found (in simulation) that the gain and time constant terms are best updated at ten times the rate at which the time delay is updated. It was also found to be necessary to limit the amount by which the gain and time constant are updated at each sample.

Representative simulation results that show the updating of the model gain are shown in Figures 7.49 to 7.52; these results, and other supplementary simulation results provided by O'Dwyer (1996j), show that convergence of the model gain to the process gain occurs, for all of the gradient algorithms, when the non-gain model parameters are equal to the corresponding process parameters. However, if the non-gain model parameters differ from the corresponding process parameters, the model gain does not converge to the process gain (unlike the behaviour of the model time delay in corresponding circumstances). The simulation conditions are as follows: $G_m e^{-s\tau_m} = 2e^{-1.4s}/(1+0.7s)$, $G_c = 1.75(1+1/0.7s)$. In Figures 7.49 and 7.50, $\tau_p = \tau_m$, $T_p = T_m$ and $K_p = 1.6$; in Figures 7.51 and 7.52, $\tau_p = \tau_m$, $T_p = T_m$ and $K_p = 2.4$.

Figure 7.49: Gain updating

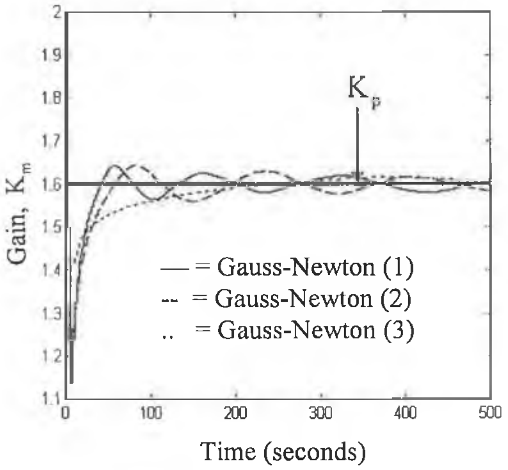


Figure 7.50: Gain updating

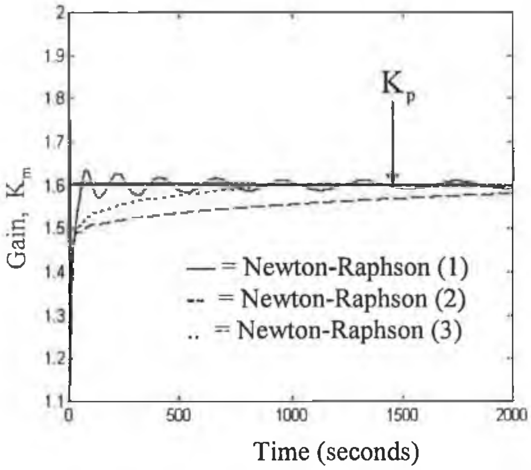


Figure 7.51: Gain updating

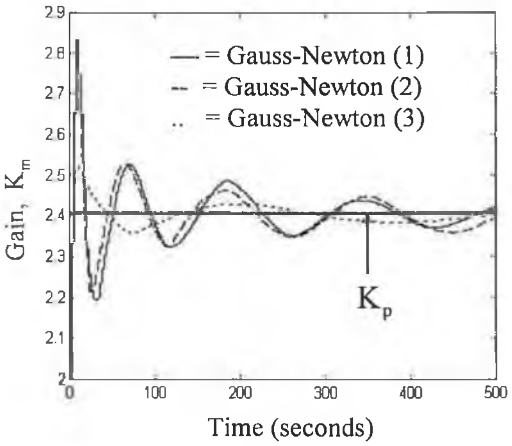
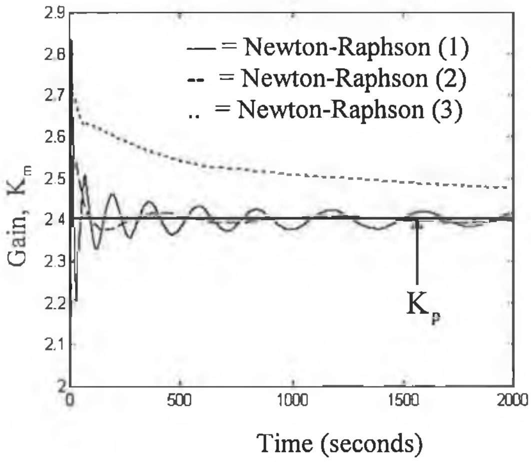


Figure 7.52: Gain updating



7.5 Parameter estimation in the modified Smith predictor

7.5.1 Introduction

The modified Smith predictor structure involves the inclusion of a dynamic element in the outer loop of the Smith predictor structure; such an arrangement facilitates an improvement in the regulator response. This issue is discussed fully in Chapter 6. In this section, the Gauss-Newton algorithms are used to separately update the gain, time constant and time delay of a FOLPD model, in such a structure.

For the modified Smith predictor structure, the transfer function, y_p^*/R, is given by equation (6.11). Simple calculations, based on Figure 6.3, show that

$$\frac{y_m^*}{R} = \frac{G_m G_c e^{-s\tau_m}}{1 + G_c G_m + G_c\left[\dfrac{1+B(s)}{1+B(s)e^{-s\tau_m}}\right]\left(G_p e^{-s\tau_p} - G_m e^{-s\tau_m}\right)} \tag{7.54}$$

with y_m^* being the model output of the modified Smith predictor. The term $(1+B(s))/(1+B(s)e^{-s\tau_m})$ is the dynamic element implemented (equation (6.9)).

7.5.2 Development of the gradient algorithms

The formulae detailed have been developed for the calculation of the sensitivity functions of the modified Smith predictor, when appropriate partial derivatives are taken of equations (6.11) and/or (7.54).

(1) Gauss-Newton (1) algorithm:

$$\frac{\partial y_p^*}{\partial \tau_p} = -s y_p^*\left[1 - \left(\frac{1+B(s)}{1+B(s)e^{-s\tau_m}}\right)\frac{y_p^*}{R}\right] \tag{7.55}$$

$$\frac{\partial y_p^*}{\partial K_p} = \frac{y_p^*}{K_p}\left[1 - \left(\frac{1+B(s)}{1+B(s)e^{-s\tau_m}}\right)\frac{y_p^*}{R}\right] \tag{7.56}$$

$$\frac{\partial y_p^*}{\partial T_p} = \frac{-sy_p^*}{\left(1+sT_p\right)}\left[1-\left(\frac{1+B(s)}{1+B(s)e^{-s\tau_m}}\right)\frac{y_p^*}{R}\right] \tag{7.57}$$

(2) <u>Gauss-Newton (2) algorithm</u>:

$$\frac{\partial e^*}{\partial \tau_p} = \frac{\partial y_p^*}{\partial \tau_p} - \frac{\partial y_m^*}{\partial \tau_p} = -sy_p^*\left[1-\left(\frac{1+B(s)}{1+B(s)e^{-s\tau_m}}\right)\left(\frac{y_p^*}{R}-\frac{y_m^*}{R}\right)\right] \tag{7.58}$$

$$\frac{\partial e^*}{\partial K_p} = \frac{\partial y_p^*}{\partial K_p} - \frac{\partial y_m^*}{\partial K_p} = \frac{y_p^*}{K_p}\left[1-\left(\frac{1+B(s)}{1+B(s)e^{-s\tau_m}}\right)\left(\frac{y_p^*}{R}-\frac{y_m^*}{R}\right)\right] \tag{7.59}$$

$$\frac{\partial e^*}{\partial T_p} = \frac{\partial y_p^*}{\partial T_p} - \frac{\partial y_m^*}{\partial T_p} = \frac{-sy_p^*}{\left(1+sT_p\right)}\left[1-\left(\frac{1+B(s)}{1+B(s)e^{-s\tau_m}}\right)\left(\frac{y_p^*}{R}-\frac{y_m^*}{R}\right)\right] \tag{7.60}$$

(3) <u>Gauss-Newton (3) algorithm</u>:

$$0.5\frac{\partial e_1^*}{\partial \tau_p} = 0.5\left(\frac{\partial y_p^*}{\partial \tau_p} + \frac{\partial y_m^*}{\partial \tau_p}\right) = -0.5sy_p^*\left[1-\left(\frac{1+B(s)}{1+B(s)e^{-s\tau_m}}\right)\left(\frac{y_p^*}{R}+\frac{y_m^*}{R}\right)\right] \tag{7.61}$$

$$0.5\frac{\partial e_1^*}{\partial K_p} = 0.5\left(\frac{\partial y_p^*}{\partial K_p} + \frac{\partial y_m^*}{\partial K_p}\right) = \frac{y_p^*}{2K_p}\left[1-\left(\frac{1+B(s)}{1+B(s)e^{-s\tau_m}}\right)\left(\frac{y_p^*}{R}+\frac{y_m^*}{R}\right)\right] \tag{7.62}$$

$$0.5\frac{\partial e_1^*}{\partial T_p} = 0.5\left(\frac{\partial y_p^*}{\partial T_p} + \frac{\partial y_m^*}{\partial T_p}\right) = \frac{-0.5sy_p^*}{\left(1+sT_p\right)}\left[1-\left(\frac{1+B(s)}{1+B(s)e^{-s\tau_m}}\right)\left(\frac{y_p^*}{R}+\frac{y_m^*}{R}\right)\right] \tag{7.63}$$

Equations (7.55) to (7.63) follow the pattern of the sensitivity function formulae developed for the Gauss-Newton (1), Gauss-Newton (2) and Gauss-Newton (3) implementations, respectively, for the Smith predictor structure (Sections 7.2.1 to 7.2.3). The formulation of these gradient algorithms for updating the model parameter values is similar to the formulations of the algorithms for a Smith predictor structure in equations (7.17), (7.24) and (7.30), respectively.

7.5.3 Parameter estimation - simulation results

The three gradient algorithms for updating the model parameters separately have been simulated individually, for the process-model combination in which both the process and model are in FOLPD form. The conditions under which the parameters are updated are similar to those in Section 7.4, with the exception that a more conservative primary controller is used to ensure control system stability over the range of process parameter values considered. Representative simulations demonstrating (a) the convergence of the model time delay to the process time delay are provided in Figures 7.55 to 7.58 (b) the convergence of the model gain to the process gain are provided in Figures 7.59 and 7.60 and (c) the convergence of the model time constant to the process time constant are provided in Figures 7.61 and 7.62. In all simulations shown, $G_m e^{-s\tau_m} = 2e^{-1.4s}/(1+0.7s)$, $G_c = 0.7(1+1/0.7s)$ and $B(s) = (0.7s+1)/(0.7s+10)$.

Time delay updating: In Figures 7.55 and 7.56, $G_p = G_m$ and τ_p = 1.2 and 1.6 seconds, respectively; in Figure 7.57, $G_p = 1.6/(1+0.5s)$, τ_p = 1.2 seconds; in Figure 7.58, $G_p = 2.4/(1+0.9s)$, τ_p = 1.6 seconds.

Figure 7.55: Time delay updating

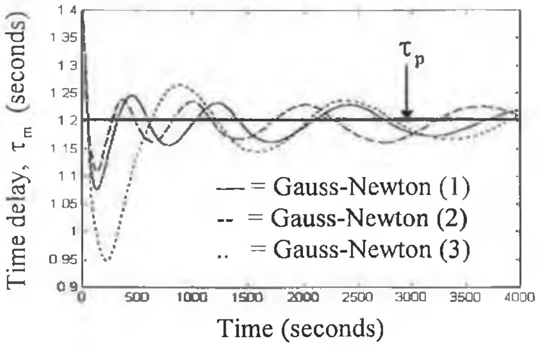


Figure 7.56: Time delay updating

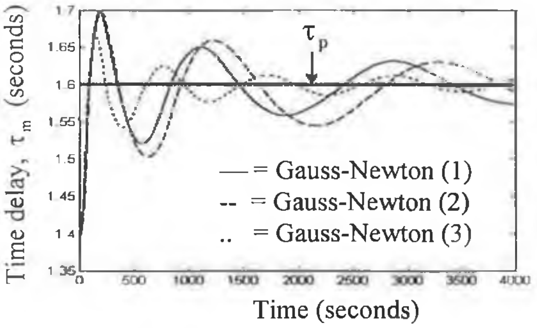


Figure 7.57: Time delay updating

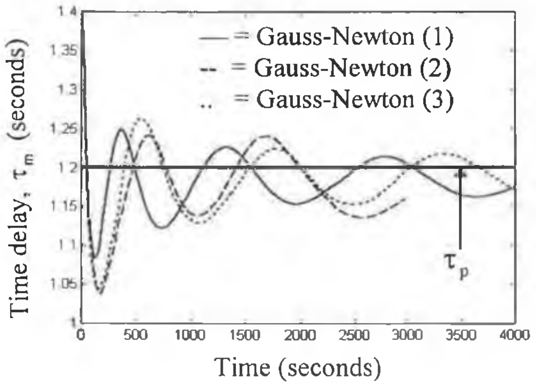
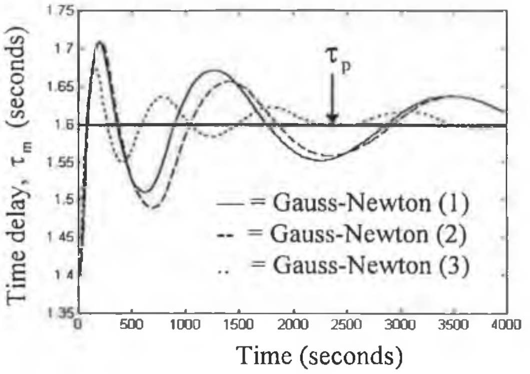


Figure 7.58: Time delay updating



Gain updating and time constant updating: In Figure 7.59, $T_p = T_m$, $\tau_p = \tau_m$ and K_p = 1.6; in Figure 7.60, $T_p = T_m$, $\tau_p = \tau_m$ and K_p = 2.4. In Figure 7.61, $K_p = K_m$, $\tau_p = \tau_m$ and T_p = 0.5 seconds; in Figure 7.62, $K_p = K_m$, $\tau_p = \tau_m$ and T_p = 0.9 seconds.

Figure 7.59: Gain updating

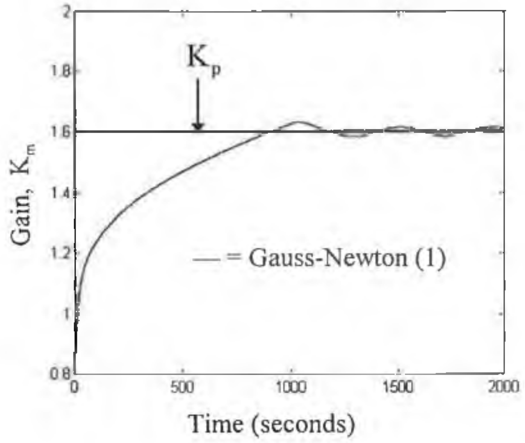


Figure 7.60: Gain updating

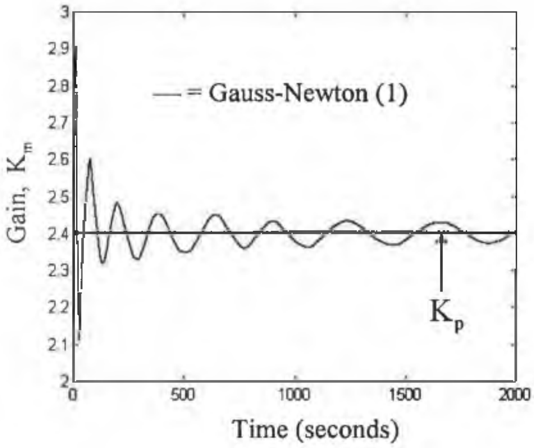


Figure 7.61: Time constant updating

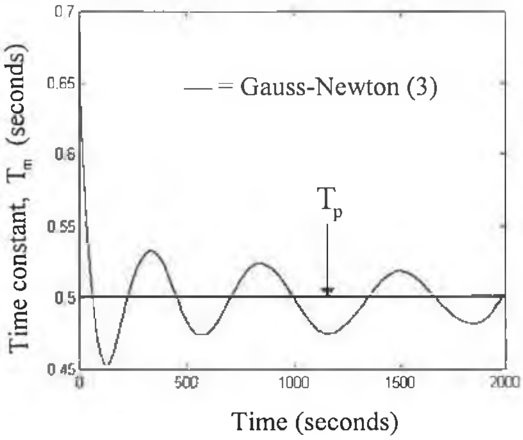
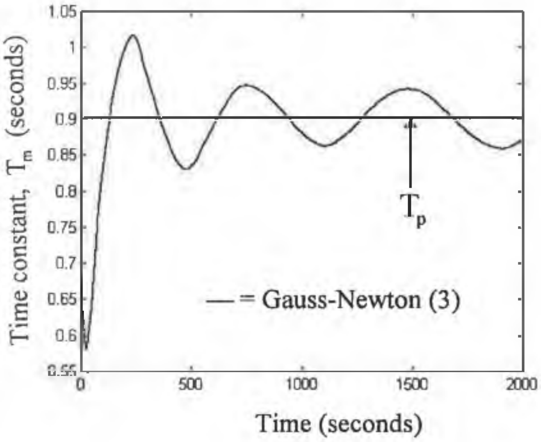


Figure 7.62: Time constant updating



These results are similar in nature to the results found when the gradient algorithms are used for model parameter updating in the Smith predictor structure. Of course, the convergence of the model parameters is slower than when the Smith predictor structure is used, because of the more aggressive primary controller used in the latter implementation. It is difficult to come to a general conclusion as to the most appropriate gradient algorithm to use for the application (as in Sections 7.4).

7.6 Analytical exploration of the algorithms used

An analytical exploration of the gradient techniques has been performed in Chapter 3 when the model parameters are estimated in open loop. The conditions under which the theorems developed in Chapter 3 are valid (such as the conditions on the process excitation signal, for example) may also apply to the closed loop gradient implementation. However, many of these theorems have been developed by assuming a first order Taylor's series expansion for the difference in time delay between the process and the model. Now, theorems will be developed without the necessity for such an approximation; the theorems developed may apply to both open loop and closed loop identification situations (provided the conditions under which the theorems are derived are fulfilled).

The analysis was performed in discrete time, for a number of process and model structures. These calculations are done in discrete time for similar reasons as reported in Chapter 3 i.e. in the discrete time domain, integer values of the process time delay appear as appropriate power terms on the numerator transfer function of the process and that a standard procedure has been defined to calculate the MSE surface, by Widrow and Stearns (1985), in the domain. The closed loop gradient algorithms are, of course, defined in continuous time; the application of the analysis performed in the discrete time domain will need to take this into account.

It is required to prove that the MSE between the process and the model output is unimodal with respect to the relevant process parameters, and is minimised when the appropriate model parameter equals the equivalent process parameter.

7.6.1 Non-delay model parameter estimation

Theorem 3.1 (Section 3.3) shows that for a first order discrete stable system, the MSE performance surface is minimised when the model gain equals the process gain and the model time constant equals the process time constant, provided the model time delay equals the process time delay and measurement noise is assumed uncorrelated to the process input and output. In addition, the input to the process and the model is assumed to be a white noise input. A corollary to this theorem is that if the process time delay index, g_p, is not equal to the model time delay index, g_m, then the

MSE function is not minimised when $K_m = K_p$ and $T_m = T_p$. A further corollary to this theorem is that the MSE function is not minimised when $K_m = K_p$ unless $g_m = g_p$ and $T_m = T_p$, and the MSE function is not minimised when $T_m = T_p$ unless $g_m = g_p$ and $K_m = K_p$. In a closed loop environment, the excitation signal to the process is not of white noise form; nevertheless, it is interesting that the simulation results in Sections 7.4.2 (Figures 7.49 to 7.54) and in Section 7.5.3 (Figures 7.59 to 7.62) show that these conclusions do apply to the closed loop identification case, provided the process input is sufficiently exciting. This is a less conservative criterion than that given in the theorem.

7.6.2 Model time delay index estimation - non-delay parameters known

Elnagger *et al.* (1990a) prove that for a first order discrete stable system of known gain and time constant, the MSE performance surface versus time delay is minimised when the model time delay index equals the process time delay index, provided the measurement noise is uncorrelated with the open loop process input. The resolution on the process time delay is assumed to be equal to one sample period. The authors also show that the MSE surface is unimodal with respect to the time delay, and that this unimodality exists for any change in the process input (such a signal is consistent with the types of signals present at the process input in closed loop applications). These conclusions conform with the simulation results in Section 7.4.1 (Figures 7.12 to 7.15) and in Section 7.5.3 (Figures 7.55 and 7.56).

7.6.3 Model time delay index estimation - non-delay parameters unknown

Elnagger *et al.* (1991) show that for a first order discrete stable system of unknown gain and time constant, the MSE performance surface versus time delay is minimised when the model time delay index equals the process time delay index. The input signal to the process is assumed to be white, though the authors state that this is a sufficient condition, rather than a necessary condition. However, the authors do not explicitly show that the MSE performance surface is unimodal with respect to the time delay, which is a requirement for the use of a gradient algorithm for time delay estimation.

It is therefore appropriate to prove that for a first order discrete stable system of unknown gain and time constant, the MSE performance surface versus time delay is minimised when the model time delay index equals the process time delay index, and that the MSE performance surface is unimodal. It will be assumed that the process excitation signal is white, which is a more rigorous requirement than that which may be achieved using a closed loop implementation; nevertheless, the proof will provide guidelines for the convergence of the parameters using the gradient algorithms in closed loop.

Theorem 7.1: For a first order discrete stable system of unknown parameters, the unimodal MSE performance surface versus time delay is minimised when the model time delay index equals the process time delay index, under the following conditions:

(a) The measurement noise is uncorrelated with the process input.

(b) The resolution on the process time delay is assumed to be equal to one sample period.

(c) The input to the process is assumed to be a white noise signal and

(d) The conditions provided in the theorem are observed on the model parameters.

Proof: The process difference equation is given by

$$y(n) = e^{-T_s/T_p} y(n-1) + K_p(1-e^{-T_s/T_p})u(n-g_p-1) + w(n) \qquad (7.64)$$

As in Chapter 3, T_s = sample period and $w(n) =$ coloured noise term. The model difference equation is (assuming the previous process output has been used in its calculation)

$$y_m(n) = e^{-T_s/T_m} y(n-1) + K_m(1-e^{-T_s/T_m})u(n-g_m-1) \qquad (7.65)$$

If the non-delay parameters are unknown, then, from equations (7.64) and (7.65), $e(n) = y(n) - y_m(n)$, is given by

$$e(n) = \left(e^{-T_s/T_p} - e^{-T_s/T_m}\right)y(n-1) + K_p(1 - e^{-T_s/T_p})u(n - g_p - 1)$$
$$-K_m(1 - e^{-T_s/T_m})u(n - g_m - 1) + w(n) \qquad (7.66)$$

The MSE performance surface, $E[e^2(n)]$, may then be calculated (from equation (7.66), using a procedure similar to that used in equations (3.22) and (3.23) (Chapter 3)), to be (O'Dwyer (1996m))

$$\left(e^{-T_s/T_p} - e^{-T_s/T_m}\right)^2 r_{yy}(0) + \left(K_p{}^2(1 - e^{-T_s/T_p})^2 + K_m{}^2(1 - e^{-T_s/T_m})^2\right)r_{uu}(0) + r_{ww}(0)$$
$$+2\left(e^{-T_s/T_p} - e^{-T_s/T_m}\right)r_{uy}(g_p) + 2\left(e^{-T_s/T_p} - e^{-T_s/T_m}\right)r_{uy}(1)$$
$$-2K_m(1 - e^{-T_s/T_m})\left[K_p(1 - e^{-T_s/T_p})r_{uu}(g_m - g_p) + (e^{-T_s/T_p} - e^{-T_s/T_m})r_{uy}(g_m)\right] \qquad (7.67)$$

assuming that the measurement noise is uncorrelated with the process input. If the excitation signal input to the process is white, then it may be shown that

$$(1)\ r_{uu}(k) = r_{uu}(0),\ \ k = 0 \text{ and } (2)\ r_{uu}(k) = 0,\ \ \text{otherwise} \qquad (7.68)$$

with

$$(3)\ r_{uy}(g_p + k) = (e^{-T_s/T_p})^{n-1} K_p(1 - e^{-T_s/T_p})r_{uu}(0),\ \ n \geq 1 \qquad (7.69)$$

and

$$(4)\ r_{uy}(g_p + n) = 0, \text{otherwise} \qquad (7.70)$$

These equations are similar to equations (3.56) to (3.58) (Chapter 3). The proof that the MSE function is unimodal with respect to the model delay, for $g_p > g_m$ and $g_p < g_m$, will be done by induction; an outline of the inductive proof (provided in full by O'Dwyer (1996m)) is as follows:

(a) For $g_p < g_m$, it may be proved, using equation (7.67), that $E\left[e^2(n)\right]$ at $g_m = g_p + 1$ > $E\left[e^2(n)\right]$ at $g_m = g_p$, provided

$$K_p(1 - e^{-T_s/T_p})r_{uu}(0) + \left(e^{-T_s/T_p} - e^{-T_s/T_m}\right)\left(r_{uy}(g_p) - r_{uy}(g_p + 1)\right) > 0 \qquad (7.71)$$

By appropriate substitution, it may be shown that this expression is true if $-(e^{-T_s/T_p} - e^{-T_s/T_m}) < 1$, which is always true.

It may also be proved, using equation (7.67), that $E[e^2(n)]$ at $g_m = g_p + n + 1 >$ $E[e^2(n)]$ at $g_m = g_p + n$, provided

$$K_p(1 - e^{-T_s/T_p})r_{uu}(n) + (e^{-T_s/T_p} - e^{-T_s/T_m})r_{uy}(g_p + n) >$$
$$K_p(1 - e^{-T_s/T_p})r_{uu}(n+1) + (e^{-T_s/T_p} - e^{-T_s/T_m})r_{uy}(g_p + n + 1) \qquad (7.72)$$

Applying equation (7.69), it may be proved by appropriate substitution that equation (7.72) is true, provided $T_p > T_m$.

(b) For $g_p > g_m$, it may be proved, by applying equations (7.67) and (7.70), that $E[e^2(n)]$ at $g_m = g_p - 1 >$ $E[e^2(n)]$ at $g_m = g_p$. It may also be proved that $E[e^2(n)]$ at $g_m = g_p - n - 1$ = $E[e^2(n)]$ at $g_m = g_p - n$, $n \neq 0$ (by applying equations (7.67) and (7.70)). Thus, the MSE performance surface is unimodal when $g_p > g_m$ only at $g_m = g_p - 1$, when the input to the process is a white noise signal. □

This theorem provides an analytical structure that helps to explain the simulation results given in Section 7.4.1 (Figures 7.20 to 7.24) and in Section 7.5.3 (Figures 7.57 and 7.58); it is interesting that these simulation results show that convergence of the model time delay to the process time delay is possible, when $K_m \neq K_p$ and $T_m \neq T_p$, even when the excitation signal to the process is not in white noise form, or when the conditions on the parameter values in the theorem are violated. This shows the conservative nature of the conclusions of the theorem.

7.6.4 Model time delay index estimation for a general model

An analytical framework on the convergence of the model time delay index, in a general model structure, will now be put in place for the case where the non-delay process and model parameters are identical. The conditions for convergence will first be calculated for a process and model in SOSPD form, as a prelude to calculating the convergence conditions for a process and model of arbitrary order. It will be demonstrated that the conditions for convergence are wider when the process and model are in SOSPD form, compared to when the process and model are of arbitrary order.

7.6.4.1 Process and model in SOSPD form

Theorem 7.2: For a second order discrete stable system of known non-delay parameters, the unimodal MSE performance surface versus time delay is minimised when the model time delay index equals the process time delay index, under the following conditions:

(a) The measurement noise is uncorrelated with the process input and

(b) The resolution on the process time delay is assumed to be equal to one sample period.

Proof: The SOSPD process difference equation is given by

$$y(n) = -a_1 y(n-1) - a_2 y(n-2) + b_1 u(n-g_p-1) + b_2 u(n-g_p-2) + w(n) \quad (7.73)$$

with a_1, a_2, b_1 and b_2 being the non-delay process parameters. The model difference equation is (assuming the previous process output has been used in its calculation)

$$y_m(n) = -a_1 y(n-1) - a_2 y(n-2) + b_1 u(n-g_m-1) + b_2 u(n-g_m-2) \quad (7.74)$$

Then, from equations (7.73) and (7.74), $e(n) = y(n) - y_m(n)$, is given by

$$e(n) = b_1\left[u(n-g_p-1) - u(n-g_m-1)\right] + b_2\left[u(n-g_p-2) - u(n-g_m-2)\right] + w(n) \quad (7.75)$$

The MSE performance surface, $E[e^2(n)]$, may then be calculated (from equation (7.75), using a procedure similar to that used in equations (3.22) and (3.23)) to be (O'Dwyer (1996m)):

$$2(b_1^2+b_2^2)r_{uu}(0)-2(b_1^2+b_2^2)r_{uu}(g_p-g_m)+4b_1b_2r_{uu}(1)$$
$$-2b_1b_2\left[r_{uu}(g_p-g_m-1)+r_{uu}(g_p-g_m+1)\right]+r_{ww}(0) \qquad (7.76)$$

assuming that the measurement noise is uncorrelated with the process input. Therefore, from equation (7.76), $E\left[e^2(n)\right]=r_{ww}(0)$ for $g_p=g_m$. The proof that the MSE function is unimodal with respect to the model delay, for $g_p>g_m$ and $g_p<g_m$, will be done by induction; an outline of the inductive proof (provided in detail by O'Dwyer (1996m)) is as follows:

(a) For $g_p>g_m$, it may be proved that $E\left[e^2(n)\right]$ at $g_m=g_p-1$ > $E\left[e^2(n)\right]$ at $g_m=g_p$, as applying equation (7.76) at $g_m=g_p-1$ gives

$$E\left[e^2(n)\right]=2\left(b_1^2+b_2^2\right)\left[r_{uu}(0)-r_{uu}(1)\right]+2b_1b_2\left[r_{uu}(1)-r_{uu}(0)\right]$$
$$+2b_1b_2\left[r_{uu}(1)-r_{uu}(2)\right]+r_{ww}(0) \qquad (7.77)$$

$$=\left(b_1-b_2\right)^2\left[r_{uu}(0)-r_{uu}(1)\right]+\left(b_1^2+b_2^2\right)\left[r_{uu}(0)-r_{uu}(1)\right]+2b_1b_2\left[r_{uu}(1)-r_{uu}(2)\right]+r_{ww}(0) \qquad (7.78)$$

This is clearly more positive than $r_{ww}(0)$.

Similarly, it may be proved that $E\left[e^2(n)\right]$ at $g_m=g_p-n-1$ > $E\left[e^2(n)\right]$ at $g_m=g_p-n$, as it may be shown by using equation (7.76) that $E\left[e^2(n)\right]$ at $g_m=g_p-n-1$ minus $E\left[e^2(n)\right]$ at $g_m=g_p-n$ equals

$$\left(b_1-b_2\right)^2\left[r_{uu}(n)-r_{uu}(n+1)\right]+\left(b_1^2+b_2^2\right)\left[r_{uu}(n)-r_{uu}(n+1)\right]+2b_1b_2\left[r_{uu}(n-1)-r_{uu}(n+2)\right] \qquad (7.79)$$

which is greater than zero.

(b) For $g_p<g_m$, it may be proved, using equation (7.76), that $E\left[e^2(n)\right]$ at $g_m=g_p+1$ > $E\left[e^2(n)\right]$ at $g_m=g_p$, as it may be shown that $E\left[e^2(n)\right]$ at $g_m=g_p+1$ equals $E\left[e^2(n)\right]$ at $g_m=g_p-1$. Similarly, it may be proved, using equation (7.76), that

$E[e^2(n)]$ at $g_m = g_p + n + 1 > E[e^2(n)]$ at $g_m = g_p + n$, as it may be shown that $E[e^2(n)]$ at $g_m = g_p + n + 1$ equals $E[e^2(n)]$ at $g_m = g_p - n - 1$ and $E[e^2(n)]$ at $g_m = g_p + n$ equals $E[e^2(n)]$ at $g_m = g_p - n$. □

The unimodality of the MSE surface versus time delay proved in this theorem exists for any change in the process input. Such a conclusion may be deduced from equations (7.78) and (7.79), as equation (7.78) equals $r_{ww}(0)$ only when the process input is a constant i.e. when the process input has a flat autocorrelation function; similarly, equation (7.79) only equals zero under the same condition. This is a significant result, as the process is in a closed loop environment; the theorem corresponds to the theorem developed by Elnagger *et al.* (1990a) for a first order discrete stable system of known non-delay parameters, described in Section 7.6.2. The conclusions reached in this theorem conform with the simulation results given in Section 7.4.1 (Figures 7.16 and 7.17).

7.6.4.2 Process and model of arbitrary order

Theorem 7.3: For an m^{th} order discrete stable system of known non-delay parameters, the unimodal MSE performance surface versus time delay is minimised when the model time delay index equals the process time delay index, under the following conditions:

(a) The measurement noise is uncorrelated with the process input.

(b) The resolution on the process time delay is assumed to be equal to one sample period and

(c) The conditions provided in the theorem are observed on the model parameters.

Proof: The process difference equation is given by

$$y(n) = \sum_{k=1}^{m} \left[-a_k y(n-k) + b_k u(n - g_p - k)\right] + w(n) \tag{7.80}$$

with $a_k, b_k, k = 1..m$ being the non-delay process parameters. The model difference equation is (assuming the previous process output has been used in its calculation)

$$y_m(n) = \sum_{k=1}^{m} \left[-a_k y(n-k) + b_k u(n-g_m-k)\right] \tag{7.81}$$

Then, from equations (7.80) and (7.81), $e(n) = y(n) - y_m(n)$, is given by

$$e(n) = \sum_{k=1}^{m}\left[b_k\{u(n-g_p-k) - u(n-g_m-k)\}\right] + w(n) \tag{7.82}$$

The MSE performance surface, $E[e^2(n)]$, may then be calculated (from equation (7.82), using a procedure similar to that used in equations (3.22) and (3.23)) to be (O'Dwyer (1996m)):

$$\begin{gathered} 2\sum_{k=1}^{m}[{b_k}^2 r_{uu}(0) - {b_k}^2 r_{uu}(g_p - g_m)] + r_{ww}(0) + \\ 2\sum_{k=1}^{m-1}\left[\sum_{j=k+1}^{m} b_k b_j\left\{2r_{uu}(j-k) - r_{uu}(g_p - g_m - j + k) - r_{uu}(g_p - g_m + j - k)\right\}\right] \end{gathered} \tag{7.83}$$

assuming that the measurement noise is uncorrelated with the process input. It may be shown, from equation (7.83), that $E\left[e^2(n)\right] = r_{ww}(0)$ for $g_p = g_m$. The proof that the MSE function is unimodal with respect to the model delay, under appropriate model parameter conditions, for $g_p > g_m$ and $g_p < g_m$, will be done by induction; an outline of the inductive proof (provided in full by O'Dwyer (1996m)) is as follows:

(a) For $g_p > g_m$, it may be proved that $E\left[e^2(n)\right]$ at $g_m = g_p - 1$ > $E\left[e^2(n)\right]$ at $g_m = g_p$, as it may be shown that, by applying equation (7.83), $E\left[e^2(n)\right]$ at $g_m = g_p - 1$ is given by

$$\begin{gathered} E\left[e^2(n)\right] = 2\sum_{k=1}^{m} {b_k}^2\left[r_{uu}(0) - r_{uu}(1)\right] + \\ \sum_{k=1}^{m} 2\sum_{k=1}^{m-1}\left[\sum_{k=j+1}^{m} b_k b_j\{2r_{uu}(j-k) - r_{uu}(k+1-j) - r_{uu}(j+1-k)\}\right] + r_{ww}(0) \end{gathered} \tag{7.84}$$

This is greater than $r_{ww}(0)$ if the excitation signal is white, for example, as equation

(7.84) then equals

$$[b_1{}^2 + b_m{}^2 + \sum_{k=2}^{m-1} (b_{k-1} - b_k)^2] r_{uu}(0) + r_{ww}(0) \tag{7.85}$$

White noise input is a sufficient, rather than a necessary condition for unimodality.

Similarly, it may be proved that $E\left[e^2(n)\right]$ at $g_m = g_p - n - 1 > E\left[e^2(n)\right]$ at $g_m = g_p - n$, using equation (7.83), provided

$$2\sum_{k=1}^{m} b_k{}^2 \left[r_{uu}(n) - r_{uu}(n+1)\right] + \\ \sum_{k=1}^{m} 2 \sum_{k=1}^{m-1} \left[\sum_{j=k+1}^{m} b_k b_j \{ r_{uu}(n-j+k) - r_{uu}(n+1-j+k) - r_{uu}(n+1+j-k) + r_{uu}(n+j-k) \} \right] > 0 \tag{7.86}$$

(b) For $g_p < g_m$, it may be proved, using equation (7.83), that $E\left[e^2(n)\right]$ at $g_m = g_p + 1$ $> E\left[e^2(n)\right]$ at $g_m = g_p$, as it may be shown that $E\left[e^2(n)\right]$ at $g_m = g_p + 1$ equals $E\left[e^2(n)\right]$ at $g_m = g_p - 1$. Similarly, it may be shown, using equation (7.83), that $E\left[e^2(n)\right]$ at $g_m = g_p + n + 1$ equals $E\left[e^2(n)\right]$ at $g_m = g_p - n - 1$ and $E\left[e^2(n)\right]$ at $g_m = g_p + n$ equals $E\left[e^2(n)\right]$ at $g_m = g_p - n$. Thus, the MSE surface is unimodal when the conditions provided in equations (7.84) and (7.86) are fulfilled. □

The conclusions reached in this theorem conform with the simulation results given in Section 7.4.1 (Figures 7.18 and 7.19).

Unfortunately, it was not possible to calculate the conditions for unimodality of the MSE function with respect to the time delay index, using the principle of induction, when both the process and the model were in general form, and when the non-delay model parameters are not equal to the non-delay process parameters.

Overall, the conclusions of the theorems conform with the appropriate simulation results quoted in Sections 7.4 and 7.5. Indeed, the results of the theorems are more conservative than many of the results achieved in simulation.

7.7 Conclusions

1. The gradient updating algorithms can facilitate a reduction in mismatch between the process and model parameter values. This reduction is mismatch has been achieved in a variety of conditions, described in the text of the chapter. The reduction in the time delay mismatch is of most interest, because mismatch in this parameter has the most influence on the stability of a Smith predictor. The most significant results are that (a) the time delay is updated, with the non-delay process and model parameters being different and (b) the time delay is updated, with the model and process structures being different. A desirable topic for further work would be to prove convergence for these two cases analytically, for an arbitrary process and model structure (though some preliminary work done in this area suggests that a proof based on the principle of induction may not be the best way of tackling such problems). A more unified approach to the problem, which would involve the application of a process order estimation strategy, as well as the time delay updating strategy, is also indicated, as the simulation results have revealed that the time delay is updated in a relatively short period of time when the model and process orders are either identical or close together; in addition, the time delay mismatch between the process and the model is only definitively reduced when the process and the model are the same order.

2. The reduction in the mismatch between the process and model time delay (in particular) allows for the possibility of updating the robust controller parameter values (in the text, the robust controller was assumed to be part of a Smith predictor structure). If an oscillatory convergence pattern is evident on the model parameters (as in the simulations taken), then this allows the calculation of the maximum mismatch from the maximum and minimum values of the parameter as it converges; this mismatch value could be used to calculate appropriate primary controller parameters (as in the procedure defined by Morari and Zafiriou (1989), for the calculation of the parameters of a robust PI primary controller). However, the methodology is by no means confined to such a controller; other possibilities include, for instance, the possibility of updating the parameters of additional dynamic terms, as well as the controller, in a modified Smith predictor structure, as the mismatch is reduced.

A further possibility is to restrict the range of variation of the parameters allowed during the updating process, so that the stability of the Smith predictor or the modified Smith predictor implementation is preserved. The implementation of such a suggestion would require that the range of variation of the process parameters is known *a priori*, though the design of the robust primary controller in the predictor structure does assume that the maximum variation of the process parameters is known *a priori*.

3. The conclusion reached in the chapter was that the Gauss-Newton (1) time delay updating algorithm is the most appropriate algorithm to use (at least for the simulations taken). This conclusion is disappointing, as the purpose of implementing the other five gradient algorithms was to try to facilitate improved performance by reducing the number of assumptions made by Marshall (1979) and Bahill (1983) in implementing the Gauss-Newton (1) algorithm. The investigation of more accurate approximations for the error (instead of those used in equations (7.2) to (7.4)) may be appropriate; a likely disadvantage is the increased complexity of such approximations.

Alternative gradient algorithms could also be employed to update the parameters. Two such alternative implementations, defined below, are based on the development used to calculate the Gauss-Newton (2) implementation and the Gauss-Newton (3) implementation, respectively. Taking these in turn, the present Gauss-Newton (2) algorithm assumes that $y_p(t,\alpha+\Delta\alpha) = y_p(t,\alpha) + \Delta\alpha\,\partial e/\partial\alpha$ (equation (7.23)) with the update values of the parameters being subsequently calculated using equation (7.24). Such an assumption is, however, unnecessary and the update values of the parameters may be more accurately calculated by solving the following relationship:

$$\begin{bmatrix} \int e\dfrac{\partial e}{\partial K_p}dt \\ \int e\dfrac{\partial e}{\partial T_p}dt \\ \int e\dfrac{\partial e}{\partial \tau_p}dt \end{bmatrix} = \begin{bmatrix} \int\left(\dfrac{\partial y_p}{\partial K_p}\right)^2 dt & \int\left(\dfrac{\partial y_p}{\partial T_p}\right)\left(\dfrac{\partial y_p}{\partial K_p}\right)dt & \int\left(\dfrac{\partial y_p}{\partial \tau_p}\right)\left(\dfrac{\partial y_p}{\partial K_p}\right)dt \\ \int\left(\dfrac{\partial y_p}{\partial T_p}\right)\left(\dfrac{\partial y_p}{\partial K_p}\right)dt & \int\left(\dfrac{\partial y_p}{\partial T_p}\right)^2 dt & \int\left(\dfrac{\partial y_p}{\partial T_p}\right)\left(\dfrac{\partial y_p}{\partial \tau_p}\right)dt \\ \int\left(\dfrac{\partial y_p}{\partial \tau_p}\right)\left(\dfrac{\partial y_p}{\partial K_p}\right)dt & \int\left(\dfrac{\partial y_p}{\partial T_p}\right)\left(\dfrac{\partial y_p}{\partial \tau_p}\right)dt & \int\left(\dfrac{\partial y_p}{\partial \tau_p}\right)^2 dt \end{bmatrix} \begin{bmatrix} \Delta K_m \\ \Delta T_m \\ \Delta \tau_m \end{bmatrix} \tag{7.87}$$

In a similar manner, the update values of the parameters in the Gauss-Newton (3) algorithm implementation are calculated using equation (7.30). This is based on implementing $\partial e/\partial\alpha \approx 0.5(\partial y_p/\partial\alpha + \partial y_m/\partial\alpha)$ (equation (7.29)). However, it is more accurate to implement $\partial e/\partial\alpha \approx \partial y_p/\partial\alpha - \partial y_m/\partial\alpha$, which means that the update values of the parameters may be calculated as follows:

$$\begin{bmatrix} \int e\dfrac{\partial e}{\partial K_p}dt \\ \int e\dfrac{\partial e}{\partial T_p}dt \\ \int e\dfrac{\partial e}{\partial \tau_p}dt \end{bmatrix} = \begin{bmatrix} \int\left(\dfrac{\partial e_1}{\partial K_p}\right)^2 dt & \int\left(\dfrac{\partial e_1}{\partial T_p}\right)\left(\dfrac{\partial e_1}{\partial K_p}\right)dt & \int\left(\dfrac{\partial e_1}{\partial \tau_p}\right)\left(\dfrac{\partial e_1}{\partial K_p}\right)dt \\ \int\left(\dfrac{\partial e_1}{\partial T_p}\right)\left(\dfrac{\partial e_1}{\partial K_p}\right)dt & \int\left(\dfrac{\partial e_1}{\partial T_p}\right)^2 dt & \int\left(\dfrac{\partial e_1}{\partial T_p}\right)\left(\dfrac{\partial e_1}{\partial \tau_p}\right)dt \\ \int\left(\dfrac{\partial e_1}{\partial \tau_p}\right)\left(\dfrac{\partial e_1}{\partial K_p}\right)dt & \int\left(\dfrac{\partial e_1}{\partial T_p}\right)\left(\dfrac{\partial e_1}{\partial \tau_p}\right)dt & \int\left(\dfrac{\partial e_1}{\partial \tau_p}\right)^2 dt \end{bmatrix} \begin{bmatrix} \Delta K_m \\ \Delta T_m \\ \Delta \tau_m \end{bmatrix} \quad (7.88)$$

Equivalent modifications could be made to the implementations of the Newton-Raphson (2) and Newton-Raphson (3) algorithms developed.

4. The limitation of the gradient based algorithms for updating the non-delay model parameters motivates the search for other algorithms to update these parameters. One such algorithm may be derived by updating the time delay using any of the gradient methods discussed in the chapter, and updating the non-delay parameters using identification algorithms (after the time delay has been updated) such as the recursive least squares (RLS) algorithm. The disadvantage of this approach is the requirement that the model and process time delays must coincide before the non-delay model parameters will converge to the corresponding process parameters. Such implementations are discussed by Kaya and Scheib (1984) and O'Connor (1989), and are outlined by O'Dwyer (1996i).

Alternatively, an analytical method to calculate the parameters of a FOLPD process, in closed loop under the control of a Smith predictor, based on the closed loop magnitude and phase information measured at a number of frequencies, could be used. Such an analytical technique has been used to calculate the parameters of a FOLPD process, in an open loop environment (Chapter 4). This method is described in more detail by O'Dwyer (1996i).

5. To summarise, the identification method is significant as the model parameters may

be updated in closed loop, using a Smith predictor structure. In broad general terms, the presence of the Smith predictor may be used to reduce the dominance of the time delay in a dominant time delay process, making the control problem more tractable.

CHAPTER 8

Conclusions

The thesis has dealt with a number of methods for estimating the process model parameters, and time delay, in both open loop and closed loop environments, together with the compensation of such systems. Two common themes exist in the different chapters of the work, the first being the use of gradient algorithms for parameter and time delay estimation and the second being the use of the Smith predictor structure for identification and control of a process with time delay.

8.1 Gradient algorithms for parameter and time delay estimation

These methods are used for process parameter and time delay estimation in open loop and in closed loop, in the time domain (Chapters 3 and 7, respectively) and in open loop, in the frequency domain (Chapter 4). The experience with these gradient methods may be compared under a number of headings:

(a) Parameter convergence. Each of the methods used facilitate convergence of model parameters to process parameters. However, interestingly, each gradient algorithm favours the convergence of some parameters over other parameters. The open loop time domain gradient algorithm used favours the convergence of the gain and time constant, when the time delay index is known (for a FOLPD process and model); convergence of the time delay index, under restrictive conditions, is only achieved when a first order Taylor's series approximation is used for the variation in time delay between the process and the model. These results are related to the conditions for unimodality of the corresponding cost function with respect to the parameter vector. The closed loop time domain gradient methods used favour the convergence of the time delay, for a wide variety of processes and models; other model parameters converge to the process parameters under very restrictive conditions (i.e. that every parameter, except the one being updated, is known). The open loop frequency domain gradient method favours the convergence of the gain and the time delay (as the corresponding cost function minimised is unimodal with respect to these parameters), with some restrictions on the convergence of the other parameters. The latter gradient method facilitates convergence over a wider range of operating conditions than the two time domain methods.

(b) Nature of the gradient algorithms used: A variety of gradient algorithms are used, in Chapters 3, 4 and 7. In the frequency domain, it is possible to formulate a gradient algorithm based on the analytical partial derivative of the cost function, with respect to the parameter vector, as the process is not parameterised in the frequency domain. In the time domain, because of the parameterisation of the process (and the subsequent dependence of the cost function, with respect to the parameter vector, on the process parameters) a gradient algorithm based on a Taylor's series expansion of the cost function about the optimum parameter vector must be used. Such an expansion will be accurate for a limited range of variation of the model parameters about the process parameters. Therefore, the implementation of the gradient algorithm in the frequency domain has an obvious advantage.

(c) Range of variation of the model parameter values: The range of variation of the model parameter values, to allow convergence to the optimum model parameter values corresponding to the global minimum of the cost function using the gradient method, may be calculated using the frequency domain approach, by calculation of the first and/or second partial derivatives of the cost function with respect to the appropriate parameter values. Indeed, as mentioned in Chapter 4, the unimodality of the cost function as the model parameters change may be checked in the same manner. It is not possible to perform such calculations in the time domain. Hence, the range of variation of the model parameters allowed, so that convergence of the model parameters to the parameters corresponding to the global minimum of the cost function is assured, for both time domain gradient methods, is not known analytically; it would have to be determined in simulation. Marshall (1979) suggests that convergence of the time delay is possible when it varies by $\pm 80\%$ about the process time delay (using the Gauss-Newton (1) algorithm, when the time delay is updated in closed loop).

(d) Choice of initial model parameter values: It is necessary that the cost function be unimodal if convergence from the initial model parameter values to the global minimum is to be achieved. A general method has not been proposed to calculate such initial values. In the time domain gradient approaches evaluated, it has been assumed that the initial values are sufficiently close to the global minimum so that unimodality exists (with the theorems in Chapters 3 and 7 providing some guidelines as to the relationships required between the process and model parameters). The estimates of the initial values could be obtained using experimental open loop or closed loop methods (as described in Chapter 2). Alternatively, a relay in series with the process in closed loop could be used to analytically calculate model parameters, from the time domain measured value produced. Such a scheme has been outlined in supplementary work by O'Dwyer (1996k), for the estimation of FOLPD model parameters. In the frequency domain approach, the initial values have been calculated analytically and, as mentioned in (c) above, unimodality of the cost function from these values may be checked. An algorithm to change the initial values used in this case, if the unimodality condition is violated, has been described in Chapter 4. In conclusion, multiple optimisation runs, each initiated at a different set of model parameters, may be required to calculate the global minimum of the cost function using the time domain gradient methods as implemented; a single optimisation run, with the possibility of appropriately changing the model parameters as a result of relevant analytical calculations may be used to calculate the global minimum of the cost function using the frequency domain gradient method as implemented.

From this discussion, and from the results contained and quoted in the thesis, it is clear that the gradient technique based in the frequency domain has significant advantages over the two time domain gradient techniques.

Some possibilities for further work using the specific gradient methods have been indicated at the conclusions of the relevant chapters, and will not be repeated here. However, a number of general recommendations are appropriate:

(1) It has already been mentioned in Chapter 3 that it may be useful to filter the data before identification, to increase the range of parameters over which the cost function is unimodal. It would be interesting to consider normalising the performance surface itself, when the time domain gradient methods are used, for similar reasons (such normalising has been applied in the implementation of the frequency domain gradient method). Pupeikis (1985) proposes one such adaptive filtering procedure to transform a multimodal cost function into unimodal form.

(2) The use of other methods, such as multiple model estimation methods or genetic algorithms, in combination with gradient methods, is indicated. A combination of estimation methods has been implemented, with some success, in the frequency domain; the use of multiple model estimation methods or genetic algorithms may

provide a means of determining the global minimum of the cost function with more certainty. The use of such methods is considered in Chapter 2.

(3) Each gradient technique has either explicitly or implicitly used a learning rate factor in the implementation of the gradient algorithm; this has been determined explicitly, in simulation, for both open loop gradient methods and included implicitly for the closed loop gradient method. The oscillatory convergence pattern of the parameters, when the closed loop gradient method is used (Chapter 7), shows that the learning rate is too large in these cases. An analytical determination of an optimum learning rate would be desirable; alternatively, the implementation of an adaptive learning rate may be appropriate (this topic has been considered briefly in Chapter 2).

8.2 The use of the Smith predictor structure for identification and control

Chapters 6 and 7 have indicated that the Smith predictor structure (and its modifications) are suitable for the identification and control of dominant time delay processes. A number of questions remain as to the practical applicability of this structure, particularly for identification; some of these issues are discussed in Section 8.1. Other issues, such as the nature of the excitation signal required at the servo input for identification (i.e. the persistent excitation condition) need further clarification. In addition, the practical application of a continuous time Smith predictor is problematic, as the (model) time delay is not easily implemented in this domain (in contrast to its straightforward implementation in the discrete time domain). Further work on these issues is recommended. In addition, it is appropriate to consider further the practical interaction of the identification and compensation strategies; one practical requirement, for instance, is that the compensated system should have a certain minimum level of robust stability and robust performance during identification. More specific recommendations for future work are described in the two relevant chapters.

8.3 Future direction of the field

It remains true to declare that the choice of identification method, and compensation method, for a process with time delay depends on the application. A number of trends in the development and application work are evident:

(1) There is still a lot of interest in the identification of FOLPD and/or SOSPD process models, using, for example, experimental closed loop methods (as detailed in Section 2.2.1.3) or by analysing the process output when a relay is switched into the closed loop compensated system in place of the controller (as detailed in Section 2.3.3.2). This appears to be due partly to the low computational intensity involved in identifying such models, and partly to concerns about how complex a model may reasonably be identified from experimental data. There appears to be scope to apply some of the methods in question to the estimation of the parameters and time delays of MIMO process models.

(2) The identification of higher order models still appears to be conditioned on the presence of *a priori* information on the process; few applications exist in which the parameters and/or time delay (of a higher order model) are identified in a black box manner from process input and output data. In addition, few unified approaches to the estimation problem have emerged; one of the rare exceptions is detailed by Chen and Zhang (1990), in which a recursive algorithm to estimate the parameters, order and time delay index of a process is described.

(3) The use of predictive controller strategies for the control of a process with time delay, such as the generalised predictive control strategy and the unified predictive control strategy, appear to be the compensation methods that are attracting increasing attention from the applications community; Kwon (1994) for instance, in a review paper, reports twenty-five such applications since 1990 in process control, mechatronics, aircraft control and medical engineering. The author also claims that "hundreds" of commercial predictive control software packages have been used in real installation examples. These are significant figures, in view of the well known (and often well founded) reluctance of applications engineers to implement controllers other than the PID controller (and its variations).

(4) Finally, developments in the applications of neural networks and genetic algorithms seem certain to have an increasing impact on the identification problem (as is suggested in Chapter 2). It also seems reasonable that neural networks, and the use of techniques such as expert systems and fuzzy logic applications will impact on the control problem. The application of robust compensation techniques to an increasing number of applications also appears inevitable; recent applications of robust strategies to the compensation of processes with time delays are discussed in Section 5.3.1.2.

9. Glossary of essential terms and symbols used

a = Term in the polynomial, B(s), used as a component of the time advance approximation (Chapter 6)

$a_{lm}, l = 0....i$ = Model denominator parameters (Chapter 7)

$a_{lm}^{'}, l = 0....i$ = Process denominator parameter estimates (for sensitivity function).. (Chapter 7)

$a_{lp}, l = 0....i$ = Process denominator parameters (Chapter 7)

$a_{xm}, x = 1...l$ = Model denominator parameters (Chapter 6)

$a_{xp}, x = 1...i$ = Process denominator parameters (Chapter 6)

α = Process parameter (Chapters 6 and 7)

α_1, α_2 = Process parameter (Chapter 7)

$\Delta\alpha$ = Process parameter change (Chapter 7)

ARMA model = Auto-Regressive Moving Average model

BPN network = BackPropagation Neural network

B(s) = Polynomial used as a component of the time advance approximation (Chapter 6)

$b_{lm}, l = 0....j$ = Model numerator parameters (Chapter 7)

$b_{lm}^{'}, l = 0....j$ = Process numerator parameter estimates (for sensitivity function).. (Chapter 7)

$b_{lp}, l = 0....j$ = Process numerator parameters (Chapter 7)

$b_{xm}, x = 1...k$ = Model numerator parameters (Chapter 6)

$b_{xp}, x = 1...j$ = Process numerator parameters (Chapter 6)

d = Time delay index (= time delay/sample time)

d(t) = Disturbance signal on the measured value

DFT = Discrete Fourier Transform

DMC algorithm = Dynamic Matrix Control algorithm

DSP = Digital Signal Processor

DTFT = Discrete Time Fourier Transform

δ = Multiplier term used in the Levenberg-Marquardt algorithm (Chapter 3)

e = Process minus model output

e_1 = Mean of process and model output (Chapter 7)

$e_1(n) = y_1(n) - y_{m1}(n)$ ... Chapter 3 (see definitions of $y_1(n)$ and $y_{m1}(n)$)

$e_2(n) = y_2(n) - y_{m2}(n)$ ... Chapter 3 (see definitions of $y_2(n)$ and $y_{m2}(n)$)

$e_3(n) = y_2(n) - y_{m3}(n)$ ... Chapter 3 (see definitions of $y_2(n)$ and $y_{m3}(n)$)

$e_4(n) = y_3(n) - y_{m4}(n)$ ... Chapter 3 (see definitions of $y_3(n)$ and $y_{m4}(n)$)

$e_5(n) = y_3(n) - y_{m5}(n)$ ... Chapter 3 (see definitions of $y_3(n)$ and $y_{m5}(n)$)

$e_6(n) = y_2(n) - y_{m6}(n)$ ... Chapter 3 (see definitions of $y_2(n)$ and $y_{m6}(n)$)

$e_7(n) = y_2(n) - y_{m7}(n)$ ... Chapter 3 (see definitions of $y_2(n)$ and $y_{m7}(n)$)

$e_8(n) = y_1(n) - y_{m8}(n)$ ... Chapter 3 (see definitions of $y_1(n)$ and $y_{m8}(n)$)

$e_9(n) = y_1(n) - y_{m9}(n)$ ... Chapter 3 (see definitions of $y_1(n)$ and $y_{m9}(n)$)

$e_{10}(n) = y_4(n) - y_{m10}(n)$ ... Chapter 3 (see definitions of $y_4(n)$ and $y_{m10}(n)$)

$e_{11}(n) = y_4(n) - y_{m11}(n)$ ... Chapter 3 (see definitions of $y_4(n)$ and $y_{m11}(n)$)

ETFE = Empirical Transfer Function Estimate

EHAC algorithm = Extended Horizon Adaptive Control algorithm

EPSAC algorithm = Extended Prediction Self-Adaptive Control algorithm

F[ ] = Fourier transform of ...

$F_1(s), F_2(s), K_1(s), K_2(s), P(s)$ = Dynamic elements in the feedback paths and the forward path of the modified Smith predictor (Chapter 6)

FFT = Fast Fourier Transform

FOL model = First Order Lag model

FOLPD model = First Order Lag Plus time Delay model

g_a = Model time delay minus model time delay index

g_b = Process time delay minus process time delay index

g_m = Model time delay index

g_p = Process time delay index

$G_c(s)$ = Controller transfer function (s domain)

$G_{L_1}(s), G_{L_2}(s)$ = Dynamic elements on the disturbance inputs (closed loop)

$G_m(s)$ = Model transfer function (s domain), not including time delay

$G_m(z)$ = Model transfer function (z domain), not including time delay

$G_p(j\omega)$ = Process frequency transfer function, at frequency ω

$|G_p(j\omega)|$ = Magnitude of the process frequency transfer function, at frequency ω

$G_p(s)$ = Process transfer function (s domain), not including time delay

$G_p(z)$ = Process transfer function (z domain), not including time delay

GMDC algorithm = Generalised MultiDelay Compensator algorithm

GPC strategy = Generalised Predictive Control strategy

I = Identity matrix

IAE criterion = Integral of Absolute Error criterion

i.i.d. random variable = Independent and identically distributed random variable

IMC strategy = Internal Model Control strategy

I-PD controller = Integral on error, Proportional and Derivative on feedback controller

ISE criterion = Integral of Squared Error criterion

ITAE criterion = Integral of Time multiplied by Absolute Error criterion

J = Cost function

K = Gain (when process gain = model gain) .... (Chapter 3)

K_2 = Time advance approximation = $(1+B(s))/(1+B(s)e^{-s\tau_m})$ .... (Chapter 6)

K_c = Proportional gain

ΔK_m = Desired change in model gain

K_m = Model gain

K_m' = Process gain estimate (sensitivity function)....(Chapter 7)

K_p = Process gain

LMS algorithm = Least Mean Squares algorithm

LQ controller = Linear Quadratic controller

LQG controller = Linear Quadratic Gaussian controller

L_1, L_2 = Disturbance inputs (closed loop system)

$\lambda(n)$ = Forgetting factor

m(t) = Manipulated variable

MAC algorithm = Model Algorithmic Control algorithm

MIMO model = Multi-Input, Multi-Output model

MISO model = Multi-Input, Single Output model

MPC algorithm = Model Predictive Control algorithm

MRAC strategy = Model Reference Adaptive Control strategy

MSE function = Mean Squared Error function

MV controller = Minimum Variance controller p = Term in the polynomial, B(s), used as a component of the time advance approximation (Chapter 6)

p.d.f. = Probability density function

P controller = Proportional controller

PCA strategy = Predictive Control Algorithm strategy

PD controller = Proportional and Derivative controller

PI controller = Proportional and Integral controller

PID controller = Proportional, Integral and Derivative controller

PIP controller = Predictive PI controller

PRBS = Pseudo-Random Binary Signal

r = Time delay variation (Chapter 3)

R = Desired input signal (to a closed loop system)

$r_{uu}(n)$ = Autocorrelation function of u(n)

$r_{uy}(n)$ = Crosscorrelation function between u(n) and y(n)

$r_{ww}(n)$ = Autocorrelation function of w(n)

$r_{yy}(n)$ = Autocorrelation function of y(n)

RELS algorithm = Recursive Extended Least Squares algorithm

RHC algorithm = Receding Horizon Control algorithm

RIV algorithm = Recursive Instrumental Variable algorithm

RLS algorithm = Recursive Least Squares algorithm

S_{α}^{T} = Sensitivity of the transfer function, T, to changes in the process parameter, α (Chapter 6)

S_m = Power spectral density of m

S_{my} = Cross power spectral density of y with respect to m

S_u = Power spectral density of u

S_{uy} = Cross power spectral density of y with respect to u

SISO model = Single Input, Single Output model

SNR = Signal to Noise Ratio

SOSPD model = Second Order System Plus time Delay model

T = Time constant (when process time constant = model time constant) .... (Chapter 3)

$T^{MSP}(s)$ = Servo transfer function of the modified Smith predictor (Chapter 6)

$T^{SP}(s)$ = Servo transfer function of the Smith predictor (Chapter 6)

T_d = Derivative time

T_i = Integral time

ΔT_m = Desired change in model time constant

T_m = Model time constant

T_m' = Process time constant estimate (sensitivity function) ... (Chapter 7)

T_p = Process time constant

T_s = Sample period

$\Delta\tau_m$ = Desired change in model time delay

τ_m = Model time delay

τ_m' = Process time delay estimate (sensitivity function) ... (Chapter 7)

τ_p = Process time delay

TDOA = Time Difference Of Arrival

Three term controller = PID controller

Time Delay Index = Time delay/sample time (integer value)

Two term controller = PI controller

u(n) = Process/Model input

u(t) = Disturbance signal on the manipulated variable

w(n) = Measurement noise term

μ = Learning rate

ω = Frequency

ω_{bw} = Bandpass filter bandwidth (Chapter 4)

ω_c = Bandpass filter centre frequency (Chapter 4)

ω_r = Frequency where the magnitude of the process is unity

ω_s = Sampling frequency

y_m = Model output

y_m^* = Model output (modified Smith predictor)

y_p = Process output

y_p^* = Process output (modified Smith predictor)

$y_1(n)$ = FOLPD process difference equation when $T_m \neq T_p$, $K_m \neq K_p$ and the process time delay is an integer multiple of the sample period (Chapter 3).

$y_2(n)$ = FOLPD process difference equation when $T_m = T_p$, $K_m = K_p$ and the process time delay is an integer multiple of the sample period (Chapter 3).

$y_3(n)$ = FOLPD process difference equation when $T_m = T_p$, $K_m = K_p$ and the process time delay is a real multiple of the sample period (Chapter 3).

$y_4(n)$ = FOLPD process difference equation when $T_m \neq T_p$, $K_m \neq K_p$ and the process time delay is a real multiple of the sample period (Chapter 3).

$y_{m1}(n)$ = FOLPD model difference equation when $T_m \neq T_p$, $K_m \neq K_p$ and the process time delay is an integer multiple of the sample period. No approximation for the time delay variation is used. The model output is a function of the previous process output (Chapter 3).

$y_{m2}(n)$ = FOLPD model difference equation when $T_m = T_p$, $K_m = K_p$ and the process time delay is an integer multiple of the sample period. A first order Taylor's series approximation for the time delay variation is used. The model output is a function of the previous process output (Chapter 3).

$y_{m3}(n)$ = FOLPD model difference equation when $T_m = T_p$, $K_m = K_p$ and the process time delay is a integer multiple of the sample period. No approximation for the time delay variation is used. The model output is a function of the previous process output (Chapter 3).

$y_{m4}(n)$ = FOLPD model difference equation when $T_m = T_p$, $K_m = K_p$ and the process time delay is an real multiple of the sample period. No approximation for the time delay variation is used. The model output is a function of the previous process output (Chapter 3).

$y_{m5}(n)$ = FOLPD model difference equation when $T_m = T_p$, $K_m = K_p$ and the process time delay is an real multiple of the sample period. A first order Taylor's series approximation for the time delay variation is used. The model output is a function of the previous process output (Chapter 3).

$y_{m6}(n)$ = FOLPD model difference equation when $T_m = T_p$, $K_m = K_p$ and the process time delay is an integer multiple of the sample period. No approximation for the time delay variation is used. The model output is a function of the previous model output (Chapter 3).

$y_{m7}(n)$ = FOLPD model difference equation when $T_m = T_p$, $K_m = K_p$ and the process time delay is an integer multiple of the sample period. A first order Taylor's series approximation for the time delay variation is used. The model output is a function of the previous model output (Chapter 3).

$y_{m8}(n)$ = FOLPD model difference equation when $T_m \neq T_p$, $K_m \neq K_p$ and the process time delay is an integer multiple of the sample period. A first order Taylor's series approximation for the time delay variation is used. The model output is a function of the previous process output (Chapter 3).

$y_{m9}(n)$ = FOLPD model difference equation when $T_m \neq T_p$, $K_m \neq K_p$ and the process time delay is an integer multiple of the sample period. A first order Taylor's series approximation for the time delay variation is used. The model output is a function of the previous model output (Chapter 3).

$y_{m10}(n)$ = FOLPD model difference equation when $T_m \neq T_p$, $K_m \neq K_p$ and the process time delay is a real multiple of the sample period. No approximation for the time delay variation is used. The model output is a function of the previous process output (Chapter 3).

$y_{m11}(n)$ = FOLPD model difference equation when $T_m \neq T_p$, $K_m \neq K_p$ and the process time delay is a real multiple of the sample period. A first order Taylor's series approximation for the time delay variation is used. The model output is a function of the previous process output (Chapter 3).

$\theta(n)$ = Parameter vector

$\phi_p(j\omega)$ = Phase of the process, at frequency ω

10. REFERENCES

Abatzoglou, T.J. (1986). A fast local maximum likelihood estimator for time delay estimation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-34**, 375-378.

Abbas, A. and Sawycr, P.E. (1995). A multiobjective design algorithm: application to the design of SISO control systems. *Computers and Chemical Engineering*, **19**, 241-248.

Abrishamkar, F. and Bekey, G.A. (1985). Estimation of time varying time delays in linear stochastic systems. *IFAC Identification and System Parameter Estimation Conference*, 1793-1798, York, U.K.

Abrishamkar, F. and Bekey, G.A. (1986). Identification of time varying delays in stochastic system via combined non-linear estimation and Viterbi algorithm. *American Control Conference*, **3**, 1619-1628, Seattle, Washington, USA.

Agamennoni, O.E., Desages, A.C. and Romagnoli, J.A. (1992). A multivariable delay compensator scheme. *Chemical Engineering Science*, **47**, 1173-1185.

Agarwal, M. and Canudas, C. (1987). On line estimation of time delay and continuous-time process parameters. *International Journal of Control*, **46**, 295-311.

Aguirre, L.A. (1992). PID tuning based on model matching. *Electronics Letters*, **28**, 2269-2271.

Alevisakis, G. and Seborg, D.E. (1974). Control of multivariable systems containing time delays using a multivariable Smith predictor. *Chemical Engineering Science*, **29**, 373-380.

Al-Chalabi, L.A. and Khalil, F.F. (1994). Dahlin minimum variance self-tuning controller. *International Journal of Control*, **60**, 747-766.

Al-Hussaini, E.K. and El-Gayaar, M.S. (1987). Robust generalised crosscorrelator filters used in time delay estimation. *IEE Proceedings, Part F*, **134**, 394-400.

Al-Hussaini, E.K. and Kassam, S.A. (1984). Robust Eckart filters for time delay estimation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-32**, 1052-1063.

Ali, W. and Burghart, J.H. (1991). Effects of lag controllers on transient response. *IEE Proceedings, Part D*, **138**, 119-122.

Allidina, A.Y., Hughes, F.M. and Tahmassebi, T. (1985). An implicit self-tuning controller for variable time-delay systems. *IFAC Identification and System Parameter Estimation Conference*, 1779-1785, York, U.K.

Al-Saggaf, U.M. (1994). An approximation technique to tune PID controllers. *Computers and Electrical Engineering*, **17,** 313-320.

Alshebeili, S.A. and Cetin, A.E. (1990). A phase reconstruction algorithm from bispectrum. *IEEE Transactions on Geoscience and Remote Sensing*, **28,** 166-170.

Alshebeili, S.A., Cetin, A.E. and Venetsanopoulos, A.N. (1991). An adaptive system identification method based on bispectrum. *IEEE Transactions on Circuits and Systems*, **38,** 967-969.

Anderson, B.D.O. and Moore, J.B. (1989). *Optimal Control Linear Quadratic Methods*, Prentice-Hall.

Antoniadis, N. and Hero, A.O. (1994). Time-delay estimation for filtered Poisson processes using an EM-type algorithm. *IEEE Transactions on Signal Processing*, **42,** 2112-2123.

Arzen, K.-E. (1987). Realisation of expert system based feedback control. *Report for Department of Automatic Control,* Lund Institute of Technology, Sweden.

Astrom, K.J. and Zhou, Z.-Y. (1981). Self-tuners with automatic adjustment of the sampling period for processes with time delays. *Report for Department of Automatic Control,* Lund Institute of Technology, Sweden.

Astrom, K.J. and Hagglund, T. (1984). Automatic tuning of simple regulators with specifications on phase and amplitude margins. *Automatica*, **20**, 645- 651.

Astrom, K.J. and Wittenmark, B. (1984). *Computer controlled systems: theory and design*, Prentice-Hall International Inc.

Astrom, K.J. and Hagglund, T. (1988). A new auto-tuning design. *Second International IFAC Symposium on the Adaptive Control of Chemical Processes,* 141-146, Lyngby, Copenhagen, Denmark.

Astrom, K.J. and Wittenmark, B. (1989). *Adaptive Control*, Addison Wesley.

Astrom, K.J. (1991). Assessment of achievable performance of simple feedback loops. *International Journal of Adaptive Control and Signal Processing*, **5**, 3-19.

Astrom, K.J., Hang, C.C., Persson, P. and Ho, W.K. (1992). Towards intelligent PID control. *Automatica*, **28**, 1-9.

Astrom, K.J., Hagglund, T. and Wallenborg, A. (1993). Automatic tuning of digital controllers with applications to HVAC plants. *Automatica*, **29**, 1333-1343.

Astrom, K.J., Hang, C.C. and Lim, B.C. (1994). A new Smith predictor for controlling a process with an integrator and long dead time. *IEEE Transactions on Automatic Control*, **39**, 343-345.

Austin, P.C., Crawford, R.A. and Carter, M.R. (1993). A generalised multivariable Smith predictor arising from new pseudo-commutativity results. *Proceedings of the IFAC 12th World Congress (Preprints)*, **2**, 223-226, Sydney, Australia.

Austin, P.C., Park, S. and Delich, J. (1994). Model-based predictive control in the process industries. *Irish DSP and Control Colloquium*, 159-168, Dublin City University.

Avitzour, D. (1991). Time delay estimation at high signal-to-noise ratio. *IEEE Transactions on Aerospace and Electronic Systems*, **27**, 234-237.

Azaria, M. and Hertz, D. (1984). Time delay estimation by generalised cross-correlation methods. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-32**, 280-285.

Azenkot, Y. and Gertner, I. (1985). The least squares estimation of time delay between two signals with unknown relative phase shift. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-33**, 308-309.

Bahill, A.T. (1983). A simple adaptive Smith-predictor for controlling time-delay systems. *IEEE Control Systems Magazine*, **3**, 2, 16-22.

Bai, E.-W. and Chyung, D.H. (1993). Improving delay estimates derived from least-squares algorithms and Pade approximations. *International Journal of Systems Science*, **24**, 745-756.

Balchen, J.G. and Lie, B. (1987). An adaptive controller based on continuous estimation of the closed loop frequency response. *Modeling, Identification and Control*, **8**, 223-240.

Banos, A. and Gomez, F. (1995). Parametric identification of transfer functions from frequency response data. *Computing and Control Engineering Journal*, **June**, 137-144.

Banyasz, Cs. and Keviczky, L. (1988). A new recursive time delay estimation method for ARMAX models. *Eighth IFAC Identification and System Parameter Estimation Symposium*, 1055-1060, Beijing, PRC.

Banyasz, Cs. and Keviczky, L. (1994). Recursive time delay estimation method. *International Journal of System Science*, **25**, 1857-1865.

Barnes, T.J.D., Wang, L. and Cleutt, W.R. (1993). A frequency domain design method for PID controllers. *Proceedings of the American Control Conference*, 890-893, San Francisco, California, USA.

Barrett, J.F. and Moir, T.J. (1986). The kepstrum method for spectral analysis. *International Journal of Control*, **43**, 29-57.

Bar-Shalom, Y., Palmieri, F., Kumar, A. and Shertukde, H.M. (1993). Analysis of wide-band cross correlation for time-delay estimation. *IEEE Transactions on Signal Processing*, **41**, 385-387.

Barthal, J.W. and Shin, Y.C. (1993). Adaptive control of nonminimum phase processes with application to the end milling process. *Proceedings of the American Control Conference*, 2449-2454, San Francisco, California, USA.

Bartlet, H., Lohmann, A.W. and Wirnitzer, B. (1984). Phase and amplitude recovery from bispectra. *Applied Optics*, **23**, 3121-3129.

Bartolini, G. and Ferrara, A. (1992). Discrete-time adaptive control of linear plants with uncertain time delay. *Proceedings of the 31st Conference on Decision and Control*, 3554-3558, Tucson, Arizona, USA.

Batur, C. (1986). Stable sub-optimum controllers for the Smith dead-time compensation. *Proceedings of the American Control Conference*, **2**, 1354-1357, Seattle, Washington, USA.

Baur, U. and Isermann, R. (1978). On-line identification of a steam heated heat exchanger with a process computer: a case study, In N.S. Rajbman (Ed.), *Identification and System Parameter Estimation*, Vol. 1, pp. 263-304. North Holland Publishing Co.

Behbehani, K., Delpasse, J.S. and Klein, K.W. (1993). A self-tuning controller with accommodation of time delay variations. *International Journal of Control*, **58,** 1215-1225.

Benouarets, M. and Atherton, D.P. (1994). Autotuning design method for a Smith predictor control scheme. *Control '94. IEE International Conference on Control*, 795-800.

Bentley, J.P. (1989). *Principles of measurement systems*, 2nd Edition, Longman Scientific and Technical.

Bethel, R.E. and Rahikka, R.G. (1987). An optimum first-order time delay tracker. *IEEE Transactions on Aerospace and Electronic Systems*, **AES-23**, 718-725.

Bethel, R.E. and Rahikka, R.G. (1990). Optimum time delay detection and tracking. *IEEE Transactions on Aerospace and Electronic Systems*, **26**, 700-711.

Bethel, R.E., Paras, G.J., Hatfalvi, E.J. and Skora, J.M. (1995). Rate-aided multisignal time delay estimation and tracking. *IEEE Transactions on Aerospace and Electronic Systems*, **31**, 1019-1041.

Bhat, N.V. and McAvoy, T.J. (1992). Determining model structures for neural models by network stripping. *Computers and Chemical Engineering*, **16**, 271-281.

Biswas, K.K. and Singh, G. (1978). Identification of stochastic time delay systems. *IEEE Transactions on Automatic Control*, **AC-23**, 504-505.

Blackowiak, A.D. and Rajan, S.D. (1995). Multi-path arrival estimates using simulated annealing: application to crosshole tomography experiment. *IEEE Journal of Oceanic Engineering*, **20**, 157-165.

Blessing, P., Baur, U. and Isermann, R. (1978). Identification of multivariable systems with recursive correlation, least squares parameter estimation and use of a compensation technique. In N.S. Rajbman (Ed.), *Identification and System Parameter Estimation*, Part 1, pp. 461-491. North Holland Publishing Co.

Bobal, V. (1995). Self tuning Ziegler-Nichols PID controller. *International Journal of Adaptive Control and Signal Processing*, **9**, 213-226.

Bogere, M.N. and Ozgen, C. (1989). On-line controller tuning of second order dead time processes. *Chemical Engineering - Research and Design*, **67**, 555-561.

Bohn, E.V. (1985). Walsh function decoupled parameter estimation equations for dynamic continuous-time models with time-delay. *IFAC Identification and System Parameter Estimation Conference*, 799-802, York, U.K.

Boje, E. and Eitelberg, E. (1991). On identification of time delay via the approximation $e^{-Ts} = (1-0.5sT)/(1+0.5sT)$. *Ninth IFAC Symposium on Identification and System Parameter Estimation*, **1**, 337-341, Budapest, Hungary.

Bokor, J. and Keviczky, L. (1984). Structure and parameter estimation of MIMO systems using elementary sub-system representation. *International Journal of Control*, **39**, 965-986.

Borg, D.A. and Giles, R.F. (1975). PID: is it an optimal regulator?. *Control Engineering*, **September**, 59-60.

Borison, U. (1979). Self-tuning regulators for a class of multivariable systems. *Automatica*, **15**, 209-215.

Boucher, R.E. and Hassab, J.C. (1981). Analysis of discrete implementation of generalized cross correlator. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-29**, 609-611.

Boudreau, D. and Kabal, P. (1992). Asymptotic receiver structures for joint maximum likelihood time delay estimation and channel identification using Gaussian signals. *IEEE Transactions on Signal Processing*, **40**, 1258-1261.

Boudreau, D. and Kabal, P. (1993). Joint time-delay estimation and adaptive recursive least squares filtering. *IEEE Transactions on Signal Processing*, **41**, 592-601.

Bradley, J.N. and Kirlin, R.L. (1984). Delay estimation by expected value. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-32**, 19-27.

Brambilla, A., Chen, S. and Scali, C. (1990). Robust tuning of conventional controllers. *Hydrocarbon Processing*, **69**, 53-58.

Brewer, D.W. (1988). Quasi-Newton methods for parameter estimation in functional differential equations. *Proceedings of the 27th IEEE Conference on Decision and Control*, **1**, 806-809, Austin, Texas, USA.

Brooks, D.H. and Nikias, C.L. (1993). The cross-bicepstrum: definition, properties and application for simultaneous reconstruction of three nonminimum phase signals. *IEEE Transactions on Signal Processing*, **41**, 2389-2404.

Brown, R.E., Maliotis, G.N. and Gibby, J.A. (1993). PID self-tuning for aluminium rolling mill. *IEEE Transactions on Industry Applications*, **29**, 578-583.

Butler, H. (1992). *Model reference adaptive control: from theory to practice*, Prentice-Hall International Series in Systems and Control Engineering.

Cabot, R.C. (1981). A note on the application of the Hilbert transform to time delay estimation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-29**, 607-609.

Camacho, E.F. and Bourdons, C. (1993). Implementation of self-tuning generalised predictive controllers for the process industry. *International Journal of Adaptive Control and Signal Processing*, **7**, 63-73.

Cameron, F. and Seborg, D.E. (1983). A self-tuning controller with a PID structure. *International Journal of Control*, **38**, 401-417.

Carter, G.C. (1981). Time delay estimation for passive sonar signal processing. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-29**, 463-470.

Carter, G.C. and Robinson, E.R. (1993). Ocean effects on time delay estimation requiring adaptation. *IEEE Journal on Oceanic Engineering*, **18**, 367-378.

Casted, E.V. (1989). Dead time and processes with stochastic disturbances. *IECON 89. 15th Annual Conference of the IEEE Industrial Electronics Society*, 452-455, Philadelphia, Pennsylvania, USA.

Chai, T.Y. (1986). A multivariable generalised self-tuning decoupling controller and application. *Proceedings of the American Control Conference*, **2**, 739-744, Seattle, USA.

Chai, T.Y. (1988). A self-tuning decoupling controller for a class of multivariable systems and global convergence analysis. *IEEE Transactions on Automatic Control*, **33**, 767-771.

Chai, T.Y. (1990). Direct adaptive decoupling control for general stochastic multivariable systems. *International Journal of Control*, **51**, 885-909.

Chai, T.-Y. and Ma, Z. (1990). A multivariable self tuning decoupling controller with consistent parameter estimates, *Proceedings of the 11th Triennial World Congress of the International Federation of Automatic Control*, **2**, 279-284, Tallinn, Estonia.

Chai, T.Y. and Wang, G. (1992). Globally convergent multivariable adaptive decoupling controller and its application to a binary distillation column. *International Journal of Control*, **55**, 415-429.

Champagne, B., Eizenmann, M. and Pasupathy, S. (1991). Exact maximum likelihood time delay estimation for short observation intervals. *IEEE Transactions on Signal Processing*, **39**, 1245-1257.

Chan, Y.T., Hattin, R.V. and Plant, J.B. (1978). The least squares estimation of time delay and its use in signal detection. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-26**, 217-222.

Chan, Y.T., Riley, J.M.F. and Plant, J.B. (1980). A parameter estimation approach to time-delay estimation and signal detection. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-28**, 8-16.

Chan, Y.T., Riley, J.M.F. and Plant, J.B. (1981). Modeling of time delay and its application to estimation of nonstationary delays. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-29**, 577-581.

Chan, Y.T. and Miskowicz, R.K. (1984). Estimation of coherence and time delay with ARMA models. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-32**, 295-303.

Chang, R.-C., Shen, S.-H. and Yu, C.-C. (1992). Derivation of transfer function from relay feedback system. *Industrial and Engineering Chemistry Research*, **31**, 855-860.

Chen, C.-L. (1989). A simple method for on-line identification and controller tuning. *AIChE Journal*, **35**, 2037-2039.

Chen, C.-L. and Jong, M.-J. (1993). Fuzzy predictive control for the time-delay system. *Second IEEE International Conference on Fuzzy Systems*, 236-240, San Francisco, California, USA.

Chen, G.P., Malik, O.P. and Hope, G.S. (1994). Generalised discrete control system design method with control limit considerations. *IEE Proceedings - Control Theory and Applications*, **141**, 39-47.

Chen, H.-F. and Guo, L. (1987). Consistent estimation of the order of stochastic control systems. *IEEE Transactions on Automatic Control*, **AC-32**, 531-535.

Chen, H.-F. and Zhang, J.-F. (1990). Identification and adaptive control for systems with unknown orders, delay and coefficients. *IEEE Transactions on Automatic Control*, **35**, 866-877.

Chen, H.-F., Guo, L. and Zhang, J.-F. (1991). Identification and adaptive control for ARMAX systems. *Lecture Notes in Control and Information Sciences*, **161**, 216-241.

Chen, R. and Loparo, K.A. (1993). Identification of time delays in linear stochastic systems. *International Journal of Control*, **57**, 1273-1291.

Cheng, F. and Venetsanopoulos, A.N. (1992). Determination of the exact phase of the bispectrum. *IEEE Transactions on Circuits and Systems - II: Analog and Digital Signal Processing*, **39**, 155-159.

Cheng, G.S. and Hung, J.C. (1985). A least-squares based self-tuning of PID controller. *Proceedings of the IEEE South East Conference*, 325-332, Raleigh, North Carolina, USA.

Cheng, Y., Karjala, T.W. and Himmelbau, D.M. (1995). Identification of nonlinear dynamic processes with unknown and variable dead time using an internal recurrent neural network. *Industrial and Engineering Chemistry Research*, **34**, 1735-1742.

Chia, T.L. and Lefkowitz, J. (1992). Robust PID tuning using IMC technology - Part 2. *Intech*, **November**, 36-39.

Chiang, H.-H. and Nikias, C.L. (1990). A new method for adaptive time delay estimation for non-Gaussian signals. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **38**, 209-219.

Chiang, R.-C. and Yu, C.-C. (1993). Monitering procedure for closed loop control: on-line identification of maximum closed-loop log modulus. *Industrial and Engineering Chemistry Research*, **32**, 90-99.

Chien, I.-L., Mellichamp, D.A. and Seborg, D.E. (1985a). Multivariable self-tuning controller with time delay compensation. *IFAC Identification and System Parameter Estimation Conference*, 1261-1266, York, U.K.

Chien, I.-L., Seborg, D.E. and Mellichamp, D.A. (1985b). A self tuning controller for systems with unknown or time varying time delays. *International Journal of Control*, **42**, 949-964.

Chien, I.-L. (1988). IMC-PID controller design - an extension. *Proceedings of the IFAC Adaptive Control of Chemical Processes Conference*, 147-152, Copenhagen, Denmark.

Chien, I.-L. and Fruehauf, P.S. (1990). Consider IMC tuning to improve controller performance. *Chemical Engineering Progress*, **October**, 33-41.

Ching, P.C. and Chan, Y.T. (1988). Adaptive time delay estimation with constraints. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **36**, 599-602.

Ching, P.C., Chan, Y.T. and Ho, K.C. (1991). Constrained adaptation for time delay estimation with multipath propagation. *IEE Proceedings, Part F*, **138**, 453-458.

Ching, P.C. and So, H.C. (1994). Two adaptive algorithms for multipath time delay estimation. *IEEE Journal on Oceanic Engineering*, **19**, 458-463.

Chisci, L. and Mosca, E. (1993). Adaptive predictive control of ARMAX plants with unknown deadtime. *IFAC Symposium on Adaptive Systems in Control and Signal Processing*, **8**, 221-226.

Chiu, K.-C., Corripio, A.B. and Smith, C.L. (1973a). Digital Control algorithms. Part I. Dahlin algorithm. *Instruments and Control Systems*, **October**, 57-59.

Chiu, K.C., Corripio, A.B. and Smith, C.L. (1973b). Digital controller algorithms. Part III. Tuning PI and PID controllers. *Instruments and Control Systems*, **December**, 41-43.

Chiu, S.-T. (1987). Estimation of the speed of a moving source from time delay estimate. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-35**, 696-697.

Chotai, A. and Young, P.C. (1985). Pole-placement design for the time delay systems using a generalised, discrete time Smith predictor. *IEE Control Conference 1985*, 218-223.

Chotai, A. and Young, P.C. (1987). Pole-placement design for time delay systems using PIP control structure. *IEE Colloquium on Control of Time Delay Systems*, **Digest No. 1987/115**, 2/1-2/7.

Chu, Y.T. and Wu, W.-T. (1986). On modelling error characterisation of multiple dead time systems. *IEEE Transactions on Automatic Control*, **AC-31**, 157-159.

Chyung, D.H. (1993). Tracking controller for output feedback linear time delay systems. *Proceedings of the American Control Conference*, 512-515, San Francisco, California, USA.

Clark, D.W. (1985). *Self-tuning and adaptive control: theory and applications.* Edited by C.J. Harris and S.A. Billings, IEE Control Engineering Series 15, Peter Peregrinus Ltd, 1985.

Clark, D.W., Mohtadi, C. and Tuffs, P.S. (1987a). Generalised predictive control - part I. The basic algorithm. *Automatica*, **23**, 137-148.

Clark, D.W., Mohtadi, C. and Tuffs, P.S. (1987b). Generalised predictive control - part II. The basic algorithm. *Automatica*, **23**, 149-160.

Clarkson, P.M. (1993). *Optimal and Adaptive Signal Processing*, CRC Press, 1993.

Clements, W.C. and Schnelle, K.B. (1963). Pulse testing for dynamic analysis. *Industrial and Engineering Chemistry Process Design and Development*, **2**, 94-102.

Clough, D.E. and Park, S.J. (1985). A novel dead-time adaptive controller. *IFAC Workshop on Adaptive Control of Chemical Processes*, 21-26, Frankfurt am Main, FRG.

Co, T.B. and Ydstie, B.E. (1990). System identification using modulating functions and fast fourier transforms. *Computers and Chemical Engineering*, **14**, 1051-1066.

Cohen, G.H. and Coon, G.A. (1953). Theoretical considerations of retarded control. *Transactions of the ASME*, **July**, 827-834.

Cook, G. and Price, M.G. (1978). Comments on 'A comparison of the Smith predictor and optimal design approaches for systems with delay in the control'. *IEEE*

Transactions on Industrial Electronics and Control Instrumentation, **IECI-25**, 180-181.

Csaki, F. and Kis, P. (1969). An engineering method for process identification. *Periodica Polytechnica, Electrical Engineering*, **13**, 73-89.

Cusani, R. (1989). Performance of fast time delay estimators. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **37**, 757-759.

Dadebo, S. and Luus, R. (1992). Optimal control of time-delay systems by dynamic programming. *Optimal Control Applications and Methods*, **13**, 29-41.

Dadebo, S.A. and McAuley, K.B. (1993). Minimum energy control of time delay systems via iterative dynamic programming. *Proceedings of the American Control Conference*, 1261-1265, San Francisco, California, USA.

Dahlin, E.B. (1968). Designing and tuning digital controllers. *Instruments and Control Systems*, **41**, 6, 77-83.

Dastych, J. (1995). Analysis and design of control loops with dead-times. *EURACO Workshop: 'Recent results in robust and adaptive control'*, 420-439, Florence, Italy.

Datta, A. and Ochoa, J. (1996). Adaptive interval model control: design and stability analysis. *Automatica*, **32**, 261-266.

David, R.A. and Stearns, S.D. (1983). Adaptive techniques for time delay estimation and tracking. *IFAC Adaptive Systems in Control and Signal Processing Conference*, 243-247, San Francisco, U.S.A.

Delopoulos, A. and Giannakis, G.B. (1994). Consistent identification of stochastic linear systems with noisy input-output data. *Automatica*, **30**, 1271-1294.

Delopoulos, A. and Giannakis, G.B. (1996). Cumulant-based identification of noisy closed loop systems. *International Journal of Adaptive Control and Signal Processing*, **10**, 303-317.

Demircioglu, H. and Gawthrop, P.J. (1991). Continuous-time generalised predictive control (CGPC). *Automatica*, **27**, 55-74.

Demuth, H. and Beale, M. (1977). *Neural networks toolbox users guide*, The MathWorks Inc.

Desbiens, A., Pomerleau, A. and Hodouin, D. (1996). Frequency based tuning of SISO controllers for two-by-two processes. *IEE Proceedings - Control Theory and Applications*, **143**, 49-56.

Deshpande, P.B. and Ash, R.H. (1983). *Elements of Computer Process Control with Advanced Control Applications*, Instrument Society of America, Prentice-Hall Inc.

Devanathan, R. (1991). Stability analysis of first order time delay systems under PID control for simultaneous parameter variation. *Control - Theory and Advanced Technology*, **7**, 531-548.

Da Silva, M.A., Gomide, F.A.C. and Amaral, W.C. (1988). A rule based procedure for selftuning PID controllers. *Proceedings of the 27th Conference on Decision and Control*, 1947-1951, Austin, Texas.

De Carvalho, J.L.M. (1993). *Dynamic Systems and Automatic Control*, Prentice-Hall.

De Keyser, R.M.C. (1986). Adaptive dead-time estimation. *IFAC Adaptive Systems in Control and Signal Processing Conference*, 385-389, Lund, Sweden.

De Paor, A.M. (1985). A modified Smith predictor and controller for unstable processes with time delay. *International Journal of Control*, **41**, 1025-1036.

De Paor, A.M. (1989). On the control of unstable time-delayed processes: approaches based on a modified Smith predictor and on a Luenberger observer. *Institute of Electrical and Electronics Engineers, UKRI Section, Control Systems Chapter*, University of Ulster at Jordanstown, Northern Ireland.

De Paor, A.M. and O'Malley, M. (1989). Controllers of Ziegler-Nichols type for unstable processes with time delay. *International Journal of Control*, **49**, 1273-1284.

De Paor, A.M. and Egan, R.P.K. (1990). Observer based sampled-data control of unstable process with time delay. *International Journal of Control*, **51**, 341-352.

De Paor, A.M. (1993). A fiftieth anniversary celebration of the Ziegler-Nichols PID controller. *International Journal of Electrical Engineering Education*, **30**, 303-316.

De Santis, R.M. (1994). A novel PID configuration for speed and position control. *Transactions of the ASME*, **116**, 542-549.

De Souza, C.E., Goodwin, G.C., Palaniswami, M. and Mayne, D.Q. (1987). Adaptive control of linear systems with variable time delay. *World Congress of the International Federation of Automatic Control*, **10**, 93-98, Munich, FRG.

De Souza, C.E., Goodwin, G.C., Mayne, D.Q. and Palaniswami, M. (1988). An adaptive control algorithm for linear systems having unknown time delay. *Automatica*, **24**, 327-341.

Dokic, M.V. and Clarkson, P.M. (1992). Real-time adaptive filters for time-delay estimation. *Mechanical Systems and Signal Processing*, **6**, 403-418.

Donoghue, J.F. (1977). A comparison of the Smith predictor and optimal design approaches for systems with delay in the controls. *IEEE Transactions on Industrial Electronics and Control Instrumentation*, **IECI-24**, 109-117.

Dorf, R.C. (1989). *Modern control systems*, 5th Edition, Addison-Wesley.

Douce, P.E. and Balmor, L. (1990). Statistics of frequency-response estimates. *IEE Proceedings, Part D*, **137**, 290-296.

Dos Santos, P.L. and De Carvalho, J.L.M. (1990). Automatic transfer function synthesis from a Bode plot. *Proceedings of the 29th Conference on Decision and Control*, 1093-1098, Honolulu, Hawaii, USA.

Draper, N.R. and Smith, H. (1981). *Applied Regression Analysis*, Second Edition, John Wiley and Sons.

Dugard, L., Goodwin, G.C. and Xianya, X. (1984). The role of the interactor matrix in multivariable stochastic adaptive control. *Automatica*, **20**, 701-709.

Dumont, G.A., Zevros, C. and Belanger, P.R. (1985). Automatic tuning of industrial PID controllers. *Proceedings of the American Control,* **3**, 1573-1578, Boston, USA.

Dumont, G.A. (1990). Control techniques in the pulp and paper industries. In C.T. Leondes (Ed.), *Control and Dynamic Systems. Advances in Theory and Applications,* Vol. 37, pp. 65-114. Academic Press Inc.

Dumont, G.A., Elnagger, A. and Elshafei, A. (1993). Adaptive predictive control of systems with time-varying time delay. *International Journal of Adaptive Control and Signal Processing*, **7**, 91-101.

Durbin, L.D. (1984a). Simulation-study of adaptive first order with deadtime compensators. *Proceedings of the American Control Conference*, **1**, 543-548.

Durbin, L.D. (1984b). Adaptive first order-plus-deadtime compensators. *IASTED Conference on Applied Identification, Modelling and Simulation (AIMS 84)*, 193-196, New Orleans, USA.

Durbin, L.D. (1985). Deadtime approximations with adaptive deadtime compensation. *Proceedings of the American Control Conference*, **3**, 1707-1712.

During, C. and Janson, K. (1993). Implementation of a recursive correlation algorithm for real-time measurement of a varying time delay. *International Conference on Signal Processing Applications and Techniques (ICSPAT '93),* 962-966, Santa Clara, California, USA.

Duttweiler, D.L. (1983). Subsampling to estimate delay with application to echo cancelling. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-31**, 1090-1099.

Edgar, T.F., Heeb, R. and Hougen, J.O. (1981). Computer aided process control system design using interactive graphics. *Computers and Chemical Engineering*, **5**, 225-232.

Egan, G. (1994). Microcomputer on-off control of temperature. Private communication.

Eitelberg, E. (1987). A regulating and tracking PI(D) controller. *International Journal of Control*, **45**, 91-95.

El-Bagoury, M.A. and Bayoumi, M.M. (1980). Multivariable self tuning augmented regulator. *Proceedings of an International Symposium on Methods and Applications in Adaptive Control*, 41-53, Bochum, Germany.

El-Hawary, F. and Mbamalu, G.A.N. (1993). Fair and Andrews's weighting-based IRWLS algorithms for time-delay estimation in underwater target tracking. *IEEE Journal of Oceanic Engineering*, **18**, 142-150.

Elnagger, A., Dumont, G.A. and Elshafei, A.-L. (1989). Recursive estimation for system of unknown delay. *Proceedings of the 28th IEEE Conference on Decision and Control*, 1809-1810, Tampa, Florida, USA.

Elnagger, A., Dumont, G.A. and Elshafei, A.-L. (1990a). System identification and adaptive control based on a variable regression for systems having unknown delay. *Proceedings of the 29th IEEE Conference on Decision and Control*, **3**, 1445-1450, Honolulu, Hawaii, USA.

Elnagger, A., Dumont, G.A. and Elshafei, A.-L. (1990b). New method for delay estimation. *Proceedings of the 29th IEEE Conference on Decision and Control*, **3**, 1629-1630, Honolulu, Hawaii, USA.

Elnagger, A., Dumont, G.A. and Elshafei, A.-L. (1991). Delay estimation using variable regression. *Proceedings of the American Control Conference*, 2812-2817, Mass., USA.

Elnagger, A., Dumont, G.A. and Elshafei, A.-L. (1992). Adaptive control with direct delay estimation. *Proceedings Control Systems '92*, 13-17, Whistler, USA.

Elnagger, A., Dumont, G.A. and Elshafei, A-L. (1993). Adaptive control with direct delay estimation. *Proceedings of the 12th IFAC World Congress (Preprints)*, **9**, 19-23, Sydney, Australia.

Eskinat, E., Johnson, S.H. and Luyben, W.L. (1993). Use of auxiliary information in system identification. *Industrial and Engineering Chemistry Research*, **32**, 1981-1992.

Etter, D.M. and Stearns, S.D. (1981). Adaptive estimation of time delays in sampled systems. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-29**, 582-587.

Faure, F. and Evans, F.J. (1969). Identification of process delay time. *IEEE Transactions on Automatic Control*, **AC-14**, 421-422.

Feder, M. and Weinstein, E. (1988). Parameter estimation of superimposed signals using the EM algorithm. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **36**, 477-489.

Feintuch, P.L., Bershad, N.J. and Reed, F.A. (1981). Time delay estimation using the LMS adaptive filter - dynamic behaviour. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-29**, 571-576.

Feng, W. (1991). On practical stability of linear multivariable feedback systems with time-delays. *Automatica*, **27**, 389-394.

Fernandes, J.M. and Ferriers, A.R. (1994). Parameter estimation for linear systems with unknown time delay and structure preservation. *Proceedings of the Asian Control Conference*, 379-382, Tokyo, Japan.

Ferretti, G., Maffezzoni, C. and Scattolini, R. (1991). Recursive estimation of time delay in sampled systems. *Automatica*, **27**, 653-661.

Ferretti, G., Maffezzoni, C. and Scattolini, R. (1995). The recursive estimation of time delay in sampled-data control systems. In C.T. Leondes (Ed.), *Control and Dynamic Systems. Advances in Theory and Applications,* Vol. 73, pp. 159-206. Academic Press Inc.

Ferretti, G., Maffezzoni, C. and Scattolini, R. (1996). On the identifiability of the time delay with least-squares methods. *Automatica*, **32**, 449-453.

Fertner, A. and Sjolund, A. (1986). Comparison of various time delay estimation methods by computer simulation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-34**, 1329-1330.

Fisher, D.G. (1991). Process control: an overview and personal perspective. *The Canadian Journal of Chemical Engineering*, **69**, 5-26.

Fong, G., Gardner, W.A. and Schell, S.V. (1994). An algorithm for improved signal-selective time-difference estimation for cyclostationary signals. *IEEE Signal Processing Letters*, **1**, 38-40.

Foss, B.A. and Wasbo, S.O. (1994). Benchmark IFAC 93: Adaptive predictive PI-control of an unknown plant. *Automatica*, **30**, 593-598.

Friedlander, B. (1980). An ARMA modeling approach to multitarget tracking. *Proceedings of the 19th IEEE Conference on Decision and Control*, 820-825.

Friedlander, B. (1982). System identification techniques for adaptive signal processing. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-30**, 240-246.

Friedlander, B. and Porat, B. (1982). A parametric technique for time delay estimation. *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, 416-419, Paris, France.

Friedlander, B. and Porat, B. (1984). A parametric technique for time delay estimation. *IEEE Transactions on Aerospace and Electronic Systems*, **AES-20**, 729-735.

Friman, M. and Waller, K.V. (1994). Autotuning of multiloop control systems. *Industrial and Engineering Chemistry Research*, **33**, 1708-1717.

Fruehauf, P.S., Chien, I.-L. and Lauritsen, M.D. (1993). Simplified IMC-PID tuning rules. *ISA/93 Advances in Instrumentation and Control*, **48**, 1745-1766, McCormick Place, Chicago, Illinois, USA.

Fujikawa, H. and Yamada, S.-I. (1991). A design method of self-tuning Smith predictor for unknown time-delay system. *Proceedings of the 1991 International Conference on Industrial Electronics - IECON '91*, **3**, 1801-1806.

Furukawa, T. and Shimemura, E. (1983). Predictive control for systems with time delay. *International Journal of Control*, **37**, 399-412.

Gabay, E. and Merhav, S.J. (1976). Identification of linear systems with time-delay operating in a closed loop in the presence of noise. *IEEE Transactions on Automatic Control*, **AC-21**, 711-716.

Gabr, M.M. (1987). Estimation of time delay and lag relationship of quadratic systems. *International Journal of Control*, **45**, 1549-1557.

Garcia, C.E. and Morari, M. (1982). Internal Model Control. 1. A unifying review and some new results. *Industrial and Engineering Chemistry Process Design and Development*, **21**, 308-323.

Garcia, C.E. and Morari, M. (1985). Internal Model Control. 2. Design procedure for multivariable systems. *Industrial and Engineering Chemistry Process Design and Development*, **24**, 472- 484.

Garcia, C.E., Prett, D.M. and Morari, M. (1989). Model predictive control: theory and practice - a survey. *Automatica*, **25**, 335-348.

Gardner, W.A. and Chen, C.-K. (1992a). Signal-selective time-difference-of-arrival estimation for passive location of man-made signal sources in highly corruptive environments, Part I: Theory and Method. *IEEE Transactions on Signal Processing*, **40**, 1168-1184.

Gardner, W.A. and Chen, C.-K. (1992b). Signal-selective time-difference-of-arrival estimation for passive location of man-made signal sources in highly corruptive environments, Part II: Algorithms and Performance. *IEEE Transactions on Signal Processing*, **40**, 1185-1197.

Gardner, W.A. and Spooner, C.M. (1992). Comparison of autocorrelation and cross-correlation methods for signal-selective TDOA estimation. *IEEE Transactions on Signal Processing*, **40**, 2606-2608.

Gawthrop, P.J. (1977). Some interpretations of the self-tuning controller. *Proceedings of the Institute of Electrical Engineers - Control & Science*, **124**, 889-894.

Gawthrop, P.J. and Nihtila, M.T. (1985). Identification of time delays using a polynomial identification method. *System and Control Letters*, **5**, 267-271.

Gawthrop, P.J., Nihtila, M. and Rad, A.B. (1989). Recursive parameter estimation of continuous-time systems with unknown time delay. *Control - Theory and Advanced Technology*, **5**, 227-248.

Gawthrop, P.J. and Nomikos, P.E. (1990). Automatic tuning of commercial PID controllers for single loop and multiloop applications. *IEEE Control Systems Magazine*, **January**, 34-42.

Gendron, S., Perrier, M., Barrette, J., Amjad, M., Holko, A. and Legault, N. (1993). Deterministic adaptive control of SISO processes using model weighting adaptation. *International Journal of Control*, **58**, 1105-1123.

George, J.D. and Goodman, D.M. (1988). Estimating time delay and transfer function parameters using wideband transient signals. *1988 International Conference on Acoustics, Speech and Signal Processing (ICASSP 88),* **5**, 2749-2752, New York, USA.

Glentis, G.-O. A. and Kalouptsidis, N. (1992). Efficient order recursive algorithms for multichannel least squares filtering. *IEEE Transactions on Signal Processing*, **40**, 1354-1374.

Goberdhansingh, E., Wang, L. and Cluett, W.R. (1992). Robust frequency domain identification. *Chemical Engineering Science*, **47**, 1989-1999.

Goodall, R.M. (1990). The delay operator z^{-1} - inappropriate for use in recursive digital filters?. *Transactions of the Institute of Measurement and Control*, **12**, 246-250.

Goodwin, G.C., Salgado, M.E. and Middleton, R.H. (1988). Indirect adaptive control: a unified approach. *Proceedings of the American Control Conference*, 2440-2445.

Goodwin, G.C., Middleton, R.H. and Poor, H.V. (1992). High-speed digital signal processing and control. *Proceedings of the IEEE*, **80**, 240-259.

Gorecki, H., Fuska, S., Grabowski, P. and Korytowski, A. (1989). *Analysis and synthesis of time delay systems*, John Wiley and Sons.

Gradshteyn, I.S. and Ryzhik, I.M. (1980). *Tables of Integrals, Series and Products - Corrected and Enlarged Edition*, Editor: A. Jeffrey, Academic Press Inc.

Grimble, M.J. (1979). Solution of the stochastic optimal control problem in the s domain for systems with time delay. *Proceedings of the Institution of Electrical Engineers - Control and Science*, **126**, 697-703.

Grimble, M.J. (1980). The solution of finite-time optimal control problems with control time delays. *Optimal Control Applications and Methods*, **1**, 263-277.

Gu, G. and Misra, P. (1992). Identification of linear time-invariant systems from frequency-response data corrupted by bounded noise. *IEE Proceedings, Part D*, **139**, 135-140.

Guez, A. and Piovoso, M. (1991). Custom neurocontroller for a time delay process. *Proceedings of the American Control Conference*, **2**, 1592-1596.

Guidorzi, R. (1975). Canonical structures in the identification of multivariable systems. *Automatica*, **11**, 361-374.

Guidorzi, R. (1981). Invariants and canonical forms for systems structural and parametric identification. *Automatica*, **17**, 117-133.

Guillaume, P., Pintelon, R. and Schoukens, J. (1992). Parametric identification of two-port models in the frequency domain. *IEEE Transactions on Instrumentation and Measurement*, **41**, 233-239.

Guillaume, P., Pintelon, R. and Schoukens, J. (1995). Robust parametric transfer function estimation using complex logarithmic frequency response data. *IEEE Transactions on Automatic Control*, **40**, 1180-1190.

Guillaume, P., Kollar, I. and Pintelon, R. (1996). Statistical analysis of nonparametric transfer function estimates. *IEEE Transactions on Instrumentation and Measurement*, **45**, 494-500.

Guo, L., Chen, H.-F. and Zhang, J.-F. (1989). Consistent parameter estimation for linear stochastic feedback control systems (CARMA model). *Automatica*, **25**, 147-151.

Gurubasavaraj, K.H. and Brogan, W.L. (1983). Identification of time delay in multiple input, multiple output discrete systems. *Proceedings of the American Control Conference*, 1253-1254.

Ha, C., Ly, U.-L. and Vagners, J. (1993). Optimal digital control with computation time-delay: a w-synthesis method. *Proceedings of the American Control Conference*, 2626-2630, San Francisco, California, USA.

Haas, W.H. and Lindquist, C.S. (1981). A synthesis of frequency domain filters for time delay estimation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-29**, 540-548.

Habermayer, M. and Keviczky, L. (1985). Investigation of an adaptive Smith controller by simulation. *IFAC Seventh Conference on Digital Computer Applications to Process Control*, 373-377, Vienna, Austria.

Habermayer, M. (1986). Quality investigation of an adaptive Smith predictor. *Simulation of Control Systems: Selected Papers from the IFAC Symposium*, 211-215, Vienna, Austria.

Habib, R.S. and Sungoor, A.M.H. (1989). Implicit design of a predictive PID controller with computational delay effect. *15th Annual Conference of the IEEE Industrial Electronics Society (IECON '89)*, pages 198-202, Philadelphia, Pennsylvania, USA.

Haest, M., Bastin, G., Gevers, M. and Wertz, V. (1990). ESPION: An expert system for system identification. *Automatica*, **26**, 85-95.

Hagglund, T. and Astrom, K.J. (1989). An industrial adaptive PID controller. *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP 89)*, 293-298, Glasgow, Scotland.

Hagglund, T. and Astrom, K.J. (1991). Industrial adaptive controllers based on frequency response techniques. *Automatica*, **27**, 599-609.

Hagglund, T. (1992). A predictive PI controller for processes with long dead times. *IEEE Control Systems Magazine*, **February**, 57-60.

Hahn, V., Unbehauen, H. and Stoffer, R. (1982). Self-tuning control of a multivariable non-minimum phase pilot plant. *Ricerche di Automatica*, **13**, 350-365.

Hakvoort, R.G. and Van den Hof, P.M.J. (1994). Frequency domain curve fitting with maximum amplitude criterion and guaranteed stability. *International Journal of Control*, **60**, 809-825.

Hammerstrom, L.G. and Waller, K.V. (1980). On optimal control of systems with delay in the control. *IEEE Transactions on Industrial Electronics and Control Instrumentation*, **IECI-27**, 301-309.

Hang, C.C. and Astrom, K.J. (1988). Practical aspects of PID auto-tuners based on relay feedback. *Proceedings of the IFAC Adaptive Control of Chemical Processes Conference*, 153-158, Copenhagen, Denmark.

Hang, C.C. (1989). The choice of controller zeroes. *IEEE Control Systems Magazine*, **January**, 72-75.

Hang, C.C., Astrom, K.J. and Ho, W.K. (1991). Refinements of the Ziegler-Nichols tuning formula. *IEE Proceedings, Part D*, **138**, 111-118.

Hang, C.C. and Chin, D. (1991). Reduced order process modelling in self tuning control. *Automatica*, **27**, 529-534.

Hang, C.C. and Sin, K.K. (1991). On-line auto tuning of PID controllers based on the cross correlation technique. *IEEE Transactions on Industrial Electronics*, **38**, 428-437.

Hang, C.C., Ho, W.K. and Cao, L.S. (1993). A comparison of two design methods for PID controllers. *ISA/93 Advances in Instrumentation and Control*, **48**, 959-967, McCormick Place, Chicago, Illinois, USA.

Hang, C.C., Loh, A.P. and Vasnani, V.U. (1994a). Relay feedback auto-tuning of cascade controllers. *IEEE Transactions on Control Systems Technology*, **2**, 42-45.

Hang, C.-C., Wang, Q.-G. and Cao, L.-S. (1994b). A novel self-tuning technique for Smith predictors. *Proceedings of the Asian Control Conference*, 327-330, Tokyo, Japan.

Hang, C.-C., Wang, Q.-G. and Cao, L.-S. (1995). Self-tuning Smith predictors for processes with long dead time. *International Journal of Adaptive Control and Signal Processing*, **9**, 255-270.

Haniff, C.A. (1991). Least-squares Fourier phase estimation from the modulo-2π bispectrum phase. *Journal of the Optical Society of America A*, **8**, 134-140.

Hannan, E.J. and Thompson, P.J. (1981). Delay estimation and the estimation of coherence and phase. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-29**, 485-490.

Hansen, P.D. (1983). Robust identification for a self-tuning controller. *Proceedings of the 3rd Yale Workshop on Application of Adaptive Systems Theory*, 47-50, New Haven, Conn., U.S.A.

Harris, S.L. and Mellichamp, D.A. (1980). On-line identification of process dynamics: use of multifrequency binary sequences. *Industrial and Engineering Chemistry Process Design and Development*, **19**, 166-174.

Harris, S.L. and Mellichamp, D.A. (1985). Controller tuning using optimisation to meet multiple closed loop criteria. *AIChE Journal*, **31**, 484-486.

Harriott, P. (1964). *Process control*, McGraw-Hill, New York, USA.

Hassab, J.C. and Boucher, R. (1976). A probabilistic analysis of time delay extraction by the cepstrum in stationary Gaussian noise. *IEEE Transactions on Information Theory*, **IT-22**, 444-454.

Hassab, J.C. and Boucher, R.E. (1981). Performance of the generalized cross correlator in the presence of a strong spectral peak in the signal. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-29**, 549-555.

Hassell, G.A. and Harper, R.E. (1994). Nonintrusive PID autotuning using dynamic modelling technology. *IEEE Conference Record of the Annual Pulp and Paper Industry Technical Conference*, 1-6, 1994.

Hayes, M.J.M. and Holohan, A. (1996). High performance control of poorly modelled systems using PID feedback. *Irish Digital Signal Processing and Control Conference (IDSPCC)*, 365-371, Dublin University, Dublin 2, Ireland.

Healey, M. (1987). *Principles of automatic control*, Hodder and Stoughton.

Hemerly, E.M. (1991). PC-based packages for identification, optimization, and adaptive control. *IEEE Control Systems Magazine*, **February**, pages 37-43.

Hertz, D. and Reiss, J. (1982). An explicit estimate of time delay between two signals, with an unknown relative phase shift. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-30**, 1006-1007.

Hertz, D. (1986). Time delay estimation by combining efficient algorithms and generalized cross correlation methods. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-34**, 1-7.

Hill, A.G. and Venable, S.W. (1989). The effect of model error on optimum PID controller tuning. *Proceedings of the ISA/89 International Conference and Exhibition*, **1**, 51-64, Philadelphia, Pa., USA.

Hinde, R.F. Jr. and Cooper, D.J. (1994). A pattern based approach to excitation diagnosis for adaptive process control. *Chemical Engineering Science*, **49**, 1403-1415.

Hinde, R.F. Jr. and Cooper, D.J. (1995). Unified excitation and performance diagnostic adaptive control framework. *AIChE Journal*, **41**, 110-112.

Hinich, M.J. and Wilson, G.R. (1992). Time delay estimation using the cross bispectrum. *IEEE Transactions on Signal Processing*, **40**, 106-113.

Hippe, P., Wurmthaler, C. and Dittrich, F. (1987). Comments on 'A regulating and tracking PI(D) controller'. *International Journal of Control*, **46**, 1851-1856.

Ho, K.C., Ching, P.C. and Chan, Y.T. (1990). Convergence speed up in adaptive time delay estimation. *International Conference on Acoustics, Speech and Signal Processing (ICASSP 90),* **3**, 1417-1420, Albuquerque, USA.

Ho, K.C., Ching, P.C. and Chan, Y.T. (1992). A new configuration for convergence speedup in adaptive time-delay estimation. *IEEE Transactions on Signal Processing*, **40**, 2683-2691.

Ho, K.C., Chan, Y.T. and Ching, P.C. (1993). Adaptive time-delay estimation in nonstationary signal and/or noise power environments. *IEEE Transactions on Automatic Control*, **41**, 2289-2299.

Ho, K.-L. and Hsu, Y.-Y. (1992). Short term load forecasting using a multilayer neural network with an adaptive learning algorithm. *IEEE Transactions on Power Systems,* **7**, 141-149.

Ho, W.K., Hang, C.C., Zhou, J.H. and Yip, C.K. (1994). Adaptive PID control of a process with underdamped response. *Proceedings of the Asian Control Conference*, 335-338, Tokyo, Japan.

Ho, W.K., Hang, C.C. and Zhou, J.H. (1995a). Performance and gain and phase margins of well-known PI tuning formulas. *IEEE Transactions on Control Systems Technology*, **3**, 245-248.

Ho, W.K., Hang, C.C. and Cao, L.S. (1995b). Tuning of PID controllers based on gain and phase margin specifications. *Automatica*, **31**, 497-502.

Hocken, R.D., Salehi, S.V. and Marshall, J.E. (1983). Time-delay mismatch and the performance of predictor control schemes. *International Journal of Control*, **38**, 433-447.

Horowitz, I. (1983). Some properties of delayed control (Smith regulator). *International Journal of Control*, **38**, 977-990.

Hsia, T.C. (1969). A discrete method for parameter identification in linear systems with transport lags. *IEEE Transactions on Aerospace and Electronic Systems*, **AES-5**, 236-239.

Hu X., Ng, T.S. and Xu H. (1988). Robust self tuner for variable time delay systems. *Eighth IFAC/IFORS Symposium on Identification and System Parameter Estimation*, 245-249, Beijing, PRC.

Huang, C.-T. and Clements, W.C. (1982). Parameter estimation for the second-order-plus-deadtime-model. *Industrial and Engineering Chemistry Process Design and Development*, **21**, 601-603.

Huang, C.-T. and Huang, M.-F. (1993). Estimation of the second order parameters from the process transient by simple calculation. *Industrial and Engineering Chemistry Research*, **32**, 228-230.

Huang, C.-T. and Chou, C.-J. (1994). Estimation of the underdamped second-order parameters from the system transient. *Industrial and Engineering Chemistry Research*, **33**, 174-176.

Huang, C.-T. and Lin, Y.-S. (1995). Tuning PID controller for open-loop unstable processes with time delay. *Chemical Engineering Communications*, **133**, 11-30.

Huang, H.-P. and Chao, Y.-C. (1982). Optimal tuning of a practical digital PID controller. *Chemical Engineering Communications*, **18**, 51-61.

Huang, H.-P., Chao, Y.-C. and Liu, P.-H. (1985). Predictive adaptive control system for unmeasured disturbances. *Industrial and Engineering Chemistry Process Design and Development*, **24**, 666-673.

Huang, H.-P., Chen, C.-L., Chao, Y.-C. and Chen P.-L. (1990). A modified Smith predictor with an approximate inverse of dead time. *AIChE Journal*, **36**, 1025-1031.

Huang, H.-P. and You, H.-S. (1994). State analytical predictor for systems with multiple time delays. *International Journal of System Science*, **25**, 991-1014, 1994.

Huang, H.-P. and You, H.-S. (1995). New robust state analytical predictor design for a general class of deadtime processes. *International Journal of Control*, **62**, 651-683.

Hunt, K.J. and Sbarbaro, D. (1991). Neural networks for nonlinear internal model control. *IEE Proceedings - Control Theory and Applications*, **138**, 431-438.

Hunt, K.J., Sbarbaro, D., Zbikowski, R. and Gawthrop, P.J. (1992). Neural networks for control systems - a survey. *Automatica*, **28**, 1083-1112.

Hwang, C. and Chuang, Y.-H. (1994). Computation of optimal reduced-order models with time delay. *Chemical Engineering Science*, **49**, 3291-3296.

Hwang, S.-H. and Chang, H.-C. (1987). A theoretical examination of closed-loop properties and tuning methods of single-loop PI controllers. *Chemical Engineering Science*, **42**, 2395-2415.

Hwang, S.-H. (1993). Adaptive dominant pole design of PID controllers based on a single closed-loop test. *Chemical Engineering Communications*, **124**, 131-152, 1993.

Hwang, S.-H. and Tseng, T.-S. (1994). Process identification and control based on dominant pole expansions. *Chemical Engineering Science*, **49**, 1973-1983.

Hwang, S.-H. and Shiu, S.-J. (1994). A new autotuning method with specifications on dominant pole placement. *International Journal of Control*, **60**, 265-282.

Hwang, S.-H. (1995). Closed-loop automatic tuning of single-input-single-output systems. *Industrial Engineering Chemistry Research*, **34**, 2406-2417.

Hwang, S.-H. and Fang, S.-M. (1995). Closed-loop tuning method based on dominant pole placement. *Chemical Engineering Communications*, **136**, 45-66.

Ichikawa, K. (1985). Frequency domain pole assignment and exact model-matching for delay systems. *International Journal of Control*, **41**, 1015-1024.

Igarashi, R., Fukushima, S., Yamada, Y., Uemara, T., Saitoh, A. and Narita, Y. (1994). Evaluation of fluctuating dead times by time interval distribution measurement. *IEEE Transactions on Instrumentation and Measurement*, **43**, 164-169.

Ikeda, M. and Ashida, T. (1979). Stabilization of linear systems with time varying delay. *IEEE Transactions on Automatic Control*, **AC-24**, 369-370.

Isermann, R., Baur, U., Bamberger, W., Kneppo, P. and Seibert, H. (1974). Comparison of six on line identification and parameter estimation methods. *Automatica*, **10**, 81-103.

Isermann, R. (1989). *Digital Control Systems Volume 1. Fundamentals, Deterministic Control*, 2nd Revised Edition, Springer-Verlag.

Isermann, R. (1991). *Digital Control Systems Volume 2. Stochastic Control, Multivariable Control, Adaptive Control, Applications*, 2nd Revised Edition, Springer-Verlag.

Isermann, R., Lachmann, K.-H. and Matko, D. (1992). *Adaptive Control Systems*, Prentice-Hall International Series in Systems and Control Engineering.

Ishida, M. and Zhan, J. (1992). Control of a process with time delay by a policy-and-experience driven neural network. *Journal of Chemical Engineering of Japan*, **25**, 763-766.

Ishida, M. and Zhan, J. (1993). Neural network control for a MIMO process with time delay. *Journal of Chemical Engineering of Japan*, **26**, 337-339.

Izzo, L., Napolitano, A. and Paura, L. (1994). Modified cyclic methods for signal selective TDOA estimation. *IEEE Transactions on Signal Processing*, **42**, 3294-3298.

Jabbari, F. (1991). Lattice filters for RLS estimation of a delta operator based model. *IEEE Transactions on Automatic Control*, **36**, 869-875.

Jacob, E.F. and Chidambaram, M. (1996). Design of controllers for unstable first-order plus time delay systems. *Computers in Chemical Engineering*, **20**, 579-584.

Jacovitti, G. and Scarano, G. (1993). Discrete time techniques for time delay estimation. *IEEE Transactions on Signal Processing*, **41**, 525-533.

Jane, R., Rix, H., Caminal, P. and Laguna, P. (1991). Alignment methods for averaging of high-resolution cardiac signals: a comparative study of performance. *IEEE Transactions on Biomedical Engineering*, **38**, 571-579.

Jerome, N.F. and Ray, W.H. (1986). High-performance multivariable control strategies for systems having time delays. *AIChE Journal*, **32**, 914-931.

Jerome, N.F. and Ray, W.H. (1992). Model-predictive control of linear multivariable systems having time delays and right-half-plane zeroes. *Chemical Engineering Science*, **47**, 763-785.

Jesus, S. and Rix, H. (1988). High resolution ECG analysis by an improved signal averaging method and comparison with a beat-to-beat approach. *Journal of Biomedical Engineering*, **10**, 25-32.

Jiang, Z.H. (1987). Block-pulse function approach to the identification of MIMO-systems and time-delay systems. *International Journal of System Science*, **18**, 1211-1220.

Jin, B. (1994). An approach for auto-tuning of PID controllers. *Proceedin* *Asian Control Conference*, 323-326, Tokyo, Japan.

Johannson, R. (1993). *System modelling and identification*, Prentice-Hall Inc.

Johnson, A. (1985). *Process dynamics, estimation and control*, Peter Peregrinus London.

Johnson, L.W. and Reiss, R.D. (1982). *Numerical Analysis*, 2nd Edition, Addiso Wesley.

Jolly, T., Bentsman, J. and Ross, D. (1993). Generalized predictive control with dynamic filtering for process control applications. *Proceedings of the American Control Conference*, 1741-1745, San Francisco, California, USA.

Jones, A.H. and Porter, B. (1985). Design of adaptive digital set-point tracking PID controllers incorporating recursive step-response matrix identifiers for multivariable plants. *Proceedings of the 24th Conference on Decision and Control*, 845-849, Fort Lauderdale, Florida, USA.

Joshi, S. and Kaufman, H. (1975). Digital adaptive controllers using second order models with transport lag. *Automatica*, **11**, 129-139.

Juricic, D. (1987). Recursive estimation of systems with time varying parameters and delays. *IFAC 10th Triennial World Congress*, 265-269, Munich, FRG.

Jussila, T.T and Koivo, H.N. (1987). Tuning of multivariable PI controllers for unknown delay-differential systems. *IEEE Transactions on Automatic Control*, **AC-32**, 364-368.

Jutan, A. and Rodriguez, E.S. (1984). Extension of a new method for on-line controller tuning. *The Canadian Journal of Chemical Engineering*, **62**, 802-807.

Jutan, A. (1989). A comparison of three closed-loop PID tuning algorithms. *AIChE Journal*, **35**, 1912-1914.

Kantor, J.C. and Andres, R.P. (1980). The analysis and design of Smith predictors using singular Nyquist arrays. *International Journal of Control*, **31**, 655-664.

Katende, E. and Jutan, A. (1993). A new constrained self-tuning PID controller. *The Canadian Journal of Chemical Engineering*, **71**, 625-633.

Kavdia, M. and Chidambaram, M. (1996). On-line controller tuning for unstable systems. *Computers in Chemical Engineering*, **20**, 301-305.

Kawabe, T. and Katayama, T. (1994). A minimax design of robust I-PD controller for system with time delay. *Proceedings of the Asian Control Conference*, 495-498, Tokyo, Japan.

Kaya, A. and Scheib, T.J. (1984). A self-tuning method for Smith predictor and PID controls. *ISA International Conference and Exhibition*, 843-855, Houston, Texas, USA.

Kaya, A. and Scheib, T.J. (1988). Tuning of PID controls of different structures. *Control Engineering*, 62-65.

Kelly, C. (1991). Time delay estimation using rational approximations to the delay operator. *B. Eng. Final Year Project Report*, Dublin City University, Dublin 9, Ireland.

Kenefic, R.J. (1981). A Bayesian approach to time delay estimation. *IEEE Transactions on Acoustics, Speech and Signal Processing,* **ASSP-29**, 611-614.

Keviczky, L. and Banyasz, Cs. (1988). A completely adaptive PID regulator. *Proceedings of the IFAC Identification and System Parameter Estimation Conference*, 89-95, Beijing, PRC.

Keviczky, L. and Banyasz, Cs. (1992). An adaptive PID regulator based on time delay estimation. *Proceedings of the 31st Conference on Decision and Control*, 3243-3248, Tucson, Arizona, USA.

Kim Y.-C., Chung C.-S. and Yang H.-S. (1987). A new algorithm for the estimation of variable time delay of discrete systems. *J. Korea Inst. Electron. Eng. (South Korea),* **24**, 52-59.

Kim, Y.H. (1995). PI controller tuning using modified relay feedback method. *Journal of Chemical Engineering of Japan*, **28**, 118-121.

Kimura, J., Nishimura, Y., Yonezawa, Y. and Saeki, M. (1994). A design method of discrete time MRACS for an unstable and nonminimum phase system with unknown time delay. *Proceedings of the Asian Control Conference*, 291-294, Tokyo, Japan.

King, M. (1984a). Taking controller tuning into the eighties. *Control and Instrumentation*, **November**, 57-59.

King, M. (1984b). Controller tuning methods revealed. *Control and Instrumentation,* **December**, 42-43.

Kinneart, M., Hanus, R. and Henrotte, J.-L. (1987). A new decoupling precompensator for indirect adaptive control of multivariable linear systems. *IEEE Transactions on Automatic Control*, **AC-32**, 455-459.

Kinneart, M. and Hanus, R. (1988). Adaptive control of multiple-input, multiple-output linear systems. *Journal A*, **29**, 9-16.

Kleinman, D.L. (1969). Optimal control of linear systems with time-delay and observation noise. *IEEE Transactions on Automatic Control*, **October**, 524-526.

Knapp, C.H. and Carter, G.C. (1976). The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-24**, 320-327.

Knapp, G.M. and Wang, H.-P. (1992). Machine fault classification: a neural network approach. *International Journal of Production Research*, **30**, 811-823.

Knobbe, E.J. (1989). Optimal control of linear stochastic systems with process and observation time delays. In C.T. Leondes (Ed.), *Control and Dynamic Systems. Advances in Theory and Applications*. Vol. 31, Part 1 of 3, Academic Press Inc.

Kocijan, J. and Korba, R. (1991). An approach to robust multivariable combustion control design. *6th Mediterranean Electro. Conf. Procs.*, **2**, 864-867, Ljubljana, Slovenia.

Koenig, D.M. (1995). Dynamic deadtime estimation for velocity dependent systems. *IEEE Transactions on Control Systems Technology*, **3**, 362-363.

Kofahl, R. and Isermann, R. (1985). A simple method for automatic tuning of PID-controllers based on process parameter estimation. *Proceedings of the American Control Conference*, 1143-1148, Boston, MA., USA.

Koganezawa, K. (1991). On-line parameter identification of non-stationary continuous system with time-variant delay. *Proceedings of the 1991 International Conference on Industrial Electronics*, **3**, 1990-1993, Kobe, Japan.

Koivo, H.N. (1980). A multivariable self-tuning regulator. *Automatica*, **16**, 351- 366.

Koivo, H.N. and Sorvari, J. (1985). On-line tuning of a multivariable PID-controller for robot manipulators. *Proceedings of the 24th IEEE Conference on Decision and Control*, 1502-1504, Ft. Lauderdale, Florida, USA.

Koivo, H., Virtanen, P. and Pusenius, M. (1988). A self-tuning controller for processes with time varying delay: application to a paper machine head box. *Proceedings of the IFAC Adaptive Control of Chemical Processes Conference*, 125-128, Denmark.

Koivo, H.N. and Tanttu, J.T. (1991). Tuning of PID Controllers: Survey of SISO and MIMO techniques. *Proceedings of the IFAC Intelligent Tuning and Adaptive Control Symposium*, 75-80, Singapore.

Kollar, I. (1992). Signal enhancement of nonsynchronised measurements for frequency domain system identification. *IEEE Transactions on Instrumentation and Measurement*, **41**, 156-159.

Kotob, S., Badran, M., Al-Atiqi, I. and Juraidan, M. (1994). Analysis and applications of self-tuning controls in a refining process: case study. *Automatica*, **30**, 1663-1675.

Krishnaswamy, P.R., Chan, B.E.M. and Rangaiah, G.P. (1987). Closed-loop tuning of process control systems. *Chemical Engineering Science*, **42**, 2173-2182.

Kristinsson, K. and Dumont, G.A. (1992). System identification and control using genetic algorithms. *IEEE Transactions on Systems, Man and Cybernetics*, **22**, 1033-1046.

Krolik, J., Eizenmann, M. and Pasupathy, S. (1988). Time delay estimation of signals with uncertain spectra. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **36**, 1801-1811.

Kumar, A. and Bar-Shalom, Y. (1993). Time-domain analysis of cross correlation for time delay estimation with an autocorrelated signal. *IEEE Transactions on Signal Processing*, **41**, 1664-1668.

Kuo, B.C. (1991). *Automatic Control Systems*, 6th Edition, Prentice-Hall Inc.

Kurz, H. (1979). Digital parameter-adaptive control of processes with unknown constant or timevarying dead time. *Proceedings of the Fifth IFAC Symposium on Identification and Parameter Estimation*, 1187-1194, Darmstadt, FRG.

Kurz, H. and Goedecke, W. (1981). Digital parameter-adaptive control of processes with unknown dead time. *Automatica*, **17**, 245-252.

Kwon, W.H. (1994). Advances in predictive control: theory and applications. *Asian Control Conference*, Tokyo, Japan.

Laguna, P., Jane, R. and Caminal, P. (1994). A time delay estimator based on the signal integral: theoretical performance and testing on ECG signals. *IEEE Transactions on Signal Processing*, **42**, 3224-3229.

Lamaire, R.O., Valavani, L., Athans, M. and Stein, G. (1991). A frequency-domain estimator for use in adaptive control systems. *Automatica*, **27**, 23-38.

Lammers, H.C. and Verbruggen, H.B. (1985). Simple self-tuning control of processes with a slowly varying time delay. *IEE Control Conference*, 393-398.

Landau, I.D. (1990). *System Identification and Control Design*, Prentice-Hall Inc.

Landau, I.D. and Voda, A. (1992). An analytical method for the auto-calibration of PID controllers. *Proceedings of the 31st Conference on Decision and Control*, 3237-3242, Tucson, Arizona, USA.

Landau, I.D. (1995). Robust digital control of systems with time delay (the Smith predictor revisited). *International Journal of Control*, **62**, 325-347.

Lang, S.-J., Gu, X.-Y. and Chai, T.-Y. (1986). A multivariable generalised self tuning controller with decoupling design. *IEEE Transactions on Automatic Control*, **AC-31**, 474-477.

Larsen, J.E. (1994). An expert system for frequency response analysis. *IEEE Control Systems magazine*, **October**, 52-57.

Latour, P.R., Koppel, L.B. and Coughanowr, D.R. (1967). Time-optimum control of chemical processes for set-point changes. *Industrial and Engineering Chemistry Process Design and Development*, **6**, 452-460.

Laughlin, D.L. and Morari, M. (1987). Smith predictor design for robust performance. *Proceedings of the American Control Conference*, 637-642, Minneapolis, Minnesota, USA.

Laughlin, D.L., Rivera, D.E. and Morari, M. (1987). Smith predictor design for robust performance. *International Journal of Control*, **46**, 477-504.

Lee, J. (1989). On-line PID controller tuning from a single, closed-loop test. *AIChE Journal*, **35**, 329-331.

Lee, J., Cho, W. and Edgar, T.F. (1990). An improved technique for PID controller tuning from closed loop tests. *AIChE Journal*, **36**, 1891-1895.

Lee, J. and Sung, S.W. (1993). Comparison of two identification methods for PID controller tuning. *AIChE Journal*, **39**, 695-697.

Lee, T.H. and Hang, C.C. (1985). A performance study of parameter estimation schemes for systems with unknown dead time. *Proceedings of the American Control Conference*, **1**, 512-516, Boston, MA., USA.

Lee, T.H., Hang, C.C., Ho, W.K. and Yue, P.K. (1993). Implementation of a knowledge-based PID autotuner. *Automatica*, **29**, 1107-1113.

Lee, T.H., Wang, Q.G., Tan, K.K. and Nungam, S. (1995a). A knowledge based approach to dead-time estimation for process control. *International Journal of Control*, **61**, 1045-1072.

Lee, T.H., Wang, Q.G. and Tan, K.K. (1995b). A modified relay-based technique for improved critical point estimation in process control. *IEEE Transactions on Control Systems Technology*, **3**, 330-337.

Lee, T.H., Wang, Q.G. and Tan, K.K. (1996). Robust Smith-predictor controller for uncertain time delay systems. *AIChE Journal*, **42**, 1033-1040.

Leffew, K.W., Stiso, M.J. and Langhorst, H. (1987). Application of digital control techniques to a laboratory extrusion process. *Proceedings of the American Control Conference*, 1002-1007, Minneapolis, Minnesota, USA.

Leigh, J.R. (1987). *Applied Control Theory*, Revised 2nd edition, Peter Peregrinus Ltd.

Leva, A. (1993). PID autotuning algorithm based on relay feedback. *IEE Proceedings, Part D*, **140**, 328-338.

Leva, A., Maffezzoni, C. and Scattolini, R. (1994). Self-tuning PI-PID regulators for stable systems with varying delay. *Automatica*, **30**, 1171-1183.

Levy, E.C. (1959). Complex curve fitting. *IRE Transactions on Automatic Control*, **AC 4**, 37-43.

Li, J., Yuan, D.C. and Jiang, C.H. (1994). A Smith predictive control structure for stochastic processes. *Proceedings of the Asian Control Conference*, 209-212, Tokyo, Japan.

Li, P.-K. (1985). The canonical form properties, structural determination and parameter estimation for multivariable system identification. *Proceedings of the IFAC Identification and System Parameter Estimation Conference*, 1665-1669, York, U.K.

Li, W., Eskinat, E. and Luyben, W.L. (1991). An improved autotune identification method. *Industrial and Engineering Chemistry Research*, **30**, 1530-1541.

Li, W. (1994). An overshootless control algorithm for large dead-time process and its application to TDM-50A series microprocessor-based multiloop controller. *Proceedings of the Asian Control Conference*, 525-528, Tokyo, Japan.

Li, Y. and Ding, Z. (1994). A new nonparametric method for linear system phase recovery from bispectrum. *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, **41**, 415-419.

Liang, G., Wilkes, D.M. and Cadzow, J.A. (1993). ARMA model order estimation based on the eigenvalues of the covariance matrix. *IEEE Transactions on Signal Processing*, **41**, 3003-3009.

Lieslehto, J., Tanttu, J.T. and Koivo, H.N. (1991). An expert system for the multivariable controller design. *Selected Papers from the IFAC Symposium on Intelligent Tuning and Adaptive Control*, 211-216, Singapore.

Lilja, M. (1988). Least squares fitting to a rational transfer function with time delay. *IEE Control Conference 1988*, 143-146.

Lilja, M. (1990). A frequency domain method for low order controller design. *Proceedings of the 11th IFAC World Congress*, **2**, 185-190, Tallinn, Estonia.

Lim, T. J. and Macleod, M.D. (1995). Adaptive algorithms for joint time delay estimation and adaptive filtering. *IEEE Transactions on Signal Processing*, **43**, 841-851.

Lin, P.L. and Wu, Y.C. (1982). Identification of multi-input multi-output linear systems from frequency response data. *Transactions of the ASME. Journal of Dynamic Systems, Measurement and Control*, **104**, 58-64.

Lin, Y.-I., Beard, J.N. and Melsheimer, S.S. (1993). Model-independent algorithms for time-optimal control of chemical processes. *AIChE Journal*, **39**, 979-988.

Litt, J. (1991). An expert system to perform on-line controller tuning. *IEEE Control Systems Magazine*, **April**, 18-23.

Liu, G. (1990). Adaptive predictor control for slowly time-varying systems with variable time delay. *Advances in Modelling and Simulation*, **20**, 9-21, 1990.

Liu, K. and Gertler, J. (1987). An intelligent adaptive scheme for biomedical control. *Proceedings of the IEEE International Symposium on Intelligent Control*, 377-382, Philadelphia, Pennsylvania, USA.

Ljung, L. (1987). *System identification: theory for the user*, Prentice-Hall Inc.

Loh, A.P. and Vasnani, V.U. (1992). Multiloop controller design for multivariable plants. *Proceedings of the 31st Conference on Decision and Control*, 181-182, Tucson, Arizona, USA.

Loh, A.P., Hang, C.C., Quek, C.K. and Vasnani, V.U. (1993). Autotuning of multiloop proportional-integral controllers using relay feedback. *Industrial and Engineering Chemistry Research*, **32**, 1102-1107.

Lourtie, I.M.G. and Moura, J.M.F. (1991). Multisource delay estimation: nonstationary signals. *IEEE Transactions on Signal Processing*, **39**, 1033-1048.

Lublinsky, B. and Fradkov, A. (1993). A comparison of adaptive and nonadaptive feedback for discrete-time plants with dead time. *IEEE Transactions on Automatic Control*, **38**, 492-495.

Lundh, M. and Astrom, K.J. (1994). Automatic initialisation of a robust self-tuning controller. *Automatica*, **30**, 1649-1662.

Luo, X., Sun, X. and Zhou, C. (1992). Time delay structure of a multiple input, multiple output system and its factorisation and compensation. *Industrial and Engineering Chemistry Research*, **31**, 1418-1421.

Luyben, W.L. (1983). *Process modeling, simulation and control for chemical engineers*, McGraw-Hill, New York.

Luyben, W.L. (1986). Simple method for tuning SISO controllers in multivariable systems. *Industrial and Engineering Chemistry Process Design and Development*, **25**, 654-660.

Lynn, P.A. and Fuerst, W. (1994). *Introductory digital signal processing with computer applications*, Revised Edition, John Wiley and Sons.

Maiti, S.N., Kapoor, N. and Saraf, D.N. (1994). Adaptive dynamic matrix control of pH. *Industrial and Engineering Chemistry Research*, **33**, 641-646.

Malik-Zafarei, M. and Jamshidi, M. (1987). *Time delay systems - analysis, optimisation and applications*. North Holland Systems and Control Series Volume 9, Elsevier Science Publishers B.V.

Mancher, H. and Hensel, H. (1985). Determination of order and deadtime for multivariable discrete-time parameter estimation methods. *Proceedings of the IFAC Identification and System Parameter Estimation Conference*, 1811-1816, York, U.K.

Manickan, T.G., Vaccaro, R.J. and Tufts, D.W. (1994). A least squares algorithm for multipath time-delay estimation. *IEEE Transactions on Signal Processing*, **42**, 3229-3233.

Manitius, A.Z. and Olbrot, A.W. (1979). Finite spectrum assignment problem for systems with delays. *IEEE Transactions on Automatic Control*, **AC-24**, 541-553.

Manitius, A.Z. (1984). Feedback controllers for a wind tunnel model involving a delay: analytical design and numerical simulation. *IEEE Transactions on Automatic Control*, **AC-29**, 1058-1068.

Mantz, R.J. and Tacconi, E.J. (1989). Complementary rules to Ziegler and Nichols' rules for a regulating and tracking controller. *International Journal of Control*, **49**, 1465-1471.

Mantz, R.J. and Tacconi, E.J. (1990). A regulating and tracking PID controller. *Industrial and Engineering Chemistry Research*, **29**, 1249-1253.

Marshall, J.E. (1979). *Control of time-delay systems*, IEE Control Engineering Series 10. Peter Peregrinus Ltd.

Marshall, J.E. (1980). Identification strategies for time delay systems. In H. Unbehauen (Ed.), *Lecture Notes in Control and Information Sciences: Methods and Applications of Adaptive Control*, **24**, Springer-Verlag, Berlin.

Marshall, J.E. and Salehi, S.V. (1982). Improvement of system performance by the use of time-delay elements. *IEE Proceedings, Part D*, **129**, 177-181.

Marshand, M. and Fu, K.-H. (1985). Frequency domain parameter estimation of aeronautical systems without and with time delay. *Proceedings of the IFAC Identification and System Parameter Estimation Conference*, 669-674, York, U.K.

Matson, C.L. (1991). Weighted-least-squares magnitude-spectrum estimation from the bispectrum. *Journal of the Optical Society of America A*, **8**, 1914-1921.

Matsuoka, T. and Ulrych, T.J. (1984). Phase estimation using the bispectrum. *Proceedings of the IEEE*, **72**, 1403-1411.

McAnany, D.E. (1993). A pole placement technique for optimum PID control parameters. *ISA/93 Advances in Instrumentation and Control Conference*, **48**, 1775-1782, McCormick Place, Chicago, Illinois, USA.

Megan, L. and Cooper, D.J. (1992). Neural network based adaptive control via temporal pattern recognition. *The Canadian Journal of Chemical Engineering*, **70**, 1209-1219.

Mendel, J.M. (1991). Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proceedings of the IEEE*, **79**, 278-305.

Menhaj, M.B. and Hagan, M.T. (1994). Analysis of delays in networked flight simulators. *IEEE Transactions on Systems, Man and Cybernetics*, **24**, 875-881.

Messer, H. and Bar-Ness, Y. (1987). Closed-loop least mean square time-delay estimator. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-35**, 413-424.

Meyer, C., Seborg, D.E. and Wood, R.K. (1976). A comparison of the Smith predictor and conventional feedback control. *Chemical Engineering Science*, **31**, 775-778.

Meyer, J.R., Whitehouse, G.D., Smith, C.L. and Murrill, P.W. (1967). Simplifying process response approximation. *Instruments and Control Systems*, **40**, 76-79.

Meyn, S.P. and Brown, L.J. (1993). Model reference adaptive control of time-varying and stochastic systems. *IEEE Transactions on Automatic Control*, **38**, 1738-1753.

Meyr, H. and Spies, G. (1984). The structure and performance of estimators for real-time estimation of randomly varying time delay. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-32**, 81-94.

Middleton, R.H. and Goodwin, G.C. (1986). Improved finite word length characteristics in digital control using delta operators. *IEEE Transactions on Automatic Control*, **AC-31**, 1015-1021.

Middleton, R.H. and Goodwin, G.C. (1990). *Digital control and estimation: a unified approach*, Prentice-Hall Inc.

Mills, P.M., Lee, P.E. and McIntosh, P. (1991). A practical study of adaptive control of an alumina calciner. *Automatica*, **27**, 441-448.

Mitchell, J.R. (1977). Comments on 'Bode compensator design'. *IEEE Transactions on Automatic Control*, **AC-22**, 869-870.

Mitchell, R.J. (1990). Multi-microprocessor control of processes with pure time delay. *Transactions of the Institute of Measurement and Control*, **12**, 58-64.

Mizuno, N. and Fujii, S. (1983). Discrete time multivariable adaptive control for non-minimum phase plants with unknown dead time. *Proceedings of the IFAC Adaptive Systems in Control and Signal Processing*, 363-368, San Francisco, USA.

Mo, L. and Bayoumi, M.M. (1993). Adaptive pole placement control of MIMO systems. *IEEE Transactions on Automatic Control*, **38**, 967-970.

Moddemeijer, R. (1989). On estimation of entropy and mutual information of continuous distributions. *Signal Processing*, **16**, 233-248.

Moore, C.F., Smith, C.L. and Murrill, P.W. (1969). Simplifying digital control dynamics for controller tuning and hardware lag effects. *Instrument Practice*, **January**, 45-49.

Morari, M. (1987). Robust process control. *Chemical Engineering Research and Design*, **65**, 462-479.

Morari, M. (1988). Advances in process control theory. *Chemical Engineering Progress*, **October**, 60-67.

Morari, M. and Zafiriou, E. (1989). *Robust process control*, Prentice-Hall Inc.

Musto, J.C. and Lauderbaugh, L.K. (1991). A heuristic search algorithm for on-line system identification. *Proceedings of the 1991 IEEE International Symposium on Intelligent Control*, 371-376, Arlington, Virginia, USA.

Nagai, K. (1986). Measurement of the time delay using the time shift property of the Discrete Fourier Transform (DFT). *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-34**, 1006-1008.

Nagarajan, R. and Sajed, K.A. (1994). Application of front end control in time-delayed plants. *Proceedings of the Institute of Mechanical Engineers, Part 1: Journal of Systems and Control Engineering*, **208**, 71-78.

Nagy, P.A.J. and Ljung, L. (1991). Estimating time-delays via state-space identification methods. *Selected papers from the 9th IFAC/IFORS Symposium on Identification and System Parameter Estimation*, 717-720, Budapest, Hungary.

Najim, K., Youlal, H. and Najim, M. (1988). Identification of a pulsed liquid-liquid extraction column. *Selected papers from the 8th IFAC/IFORS Symposium on Identification and System Parameter Estimation*, 1165-1170, Beijing, PRC.

Namazi, M. and Stuller, J.A. (1987). A new approach to signal registration with the emphasis on variable time delay estimation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-35**, 1649-1660.

Namazi, N.M. and Biswal, B. (1992). Variable time delay estimation in coloured and correlated noise environment. *IEEE Transactions on Aerospace and Electronic Systems*, **28**, 597-603.

Narendra, K.S. and Parthasarathy, K. (1990). Identification and control of dynamical systems using neural networks. *IEEE Transactions on Neural Networks*, **1**, 4-27.

Nehorai, A. and Morf, M. (1982). Estimation of time difference of arrivals for multiple ARMA sources by a pole decomposition method. *Proceedings of the 21st Conference on Decision and Control*, **3**, 1000-1002, Orlando, Florida, USA.

Newell, R.B. and Lee, P.L. (1989). *Applied process control - a case study*, Prentice-Hall.

Ng, L.C. and Bar-Shalom, Y. (1982). Time delay estimation in a multitarget environment. *Proceedings of the 21st IEEE Conference on Decision and Control*, **3**, 1135-1139, Orlando, Florida, USA.

Ng, L.C. and Bar-Shalom, Y. (1986). Multisensor multitarget time delay vector estimation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-34**, 669-677.

Nikias, C.L. and Raghuveer, M.R. (1987). Bispectrum estimation: a digital signal processing framework. *Proceedings of the IEEE*, **75**, 869-891.

Nikias, C.L. and Pan, R. (1988). Time delay estimation in unknown Gaussian spatially correlated noise. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **36**, 1706-1714.

Nikias, C.L. and Mendel, J.M. (1993). Signal processing with higher-order spectra. *IEEE Signal Processing Magazine*, **July**, 10-37.

Nikias, C.L. and Petropulu, A.P. (1993). *Higher-order spectra analysis: a nonlinear signal processing framework*, Prentice-Hall Inc.

Nishikawa, Y., Sannomiya, N., Ohta, T. and Tanaka, H. (1984). A method for auto-tuning of PID control parameters. *Automatica*, **20**, 321-332.

Niu, S., Xiao, D. and Fisher, D.G. (1990). A recursive algorithm for simultaneous estimation of model order and parameters. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **38**, 884-886.

Niu, S. and Fisher, D.G. (1994). Simultaneous structure identification and parameter estimation of multivariable systems. *International Journal of Control*, **59**, 1127-1141.

O'Connor, M. (1989). Computer based control of time delay systems using a Smith Predictor. *B.Eng. Final Year Project Report*, National Institute of Higher Education, Dublin 9, Ireland.

O'Donnell, J.D. (1991). Model order determination for system identification. *B.Eng. Final Year Project Report*, Dublin City University, Dublin 9, Ireland.

O'Dwyer, A. (1992). A comparison of techniques for time delay estimation. *M.Eng. dissertation*, Dublin City University, Dublin 9, Ireland.

O'Dwyer, A. (1993). A comparison of techniques for time delay estimation. *Irish DSP and Control Colloquium*, University College Dublin, Dublin 4, Ireland.

O'Dwyer, A. and Ringwood, J.V. (1994a). Model parameter and time delay estimation using gradient methods. *Irish DSP and Control Colloquium*, Dublin City University, Dublin 9, Ireland.

O'Dwyer, A. and Ringwood, J.V. (1994b). Estimation of the parameters of a first order lag plus deadtime system using gradient methods. *EURACO Workshop: 'Recent results in robust and adaptive control'*, Dublin University, Dublin 2, Ireland.

O'Dwyer, A. and Ringwood, J.V. (1995). Model parameter and time delay estimation in the frequency domain. *Irish DSP and Control Colloquium,* Queens University, Belfast, Northern Ireland.

O'Dwyer, A. and Ringwood, J.V. (1996). The control of a process with time delay by using a modified Smith predictor compensator. *Irish DSP and Control Conference,* Dublin University, Dublin 2, Ireland.

O'Dwyer, A. (1996a). Literature survey into process parameter and time delay estimation, for SISO and MIMO processes. *Technical Report EE/JVR/96/01*, Dublin City University, Dublin 9, Ireland.

O'Dwyer, A. (1996b). Literature survey into parameter optimised controllers, for the compensation of SISO and MIMO processes. *Technical Report EE/JVR/96/02*, Dublin City University, Dublin 9, Ireland.

O'Dwyer, A. (1996c). Literature survey into structurally optimised controllers, for the compensation of SISO and MIMO processes. *Technical Report EE/JVR/96/03*, Dublin City University, Dublin 9, Ireland.

O'Dwyer, A. (1996d). Reference list for Ph.D. work. *Technical Report EE/JVR/96/04*, Dublin City University, Dublin 9, Ireland.

O'Dwyer, A. (1996e). Open loop time domain gradient methods of parameter and time delay estimation. *Technical Report EE/JVR/96/05*, Dublin City University, Dublin 9, Ireland.

O'Dwyer, A. (1996f). Compensation of processes with time delays by using an appropriately modified Smith predictor compensator. *Technical Report EE/JVR/96/06*, Dublin City University, Dublin 9, Ireland.

O'Dwyer, A. (1996g). Frequency domain methods of parameter and time delay estimation. *Technical Report EE/JVR/96/07*, Dublin City University, Dublin 9, Ireland.

O'Dwyer, A. (1996h). Simulation results associated with frequency domain methods of parameter and time delay estimation. *Technical Report EE/JVR/96/08*, Dublin City University, Dublin 9, Ireland.

O'Dwyer, A. (1996i). Estimation of model parameters (including time delay) in a Smith predictor structure. *Technical Report EE/JVR/96/09*, Dublin City University, Dublin 9, Ireland.

O'Dwyer, A. (1996j). Simulation results associated with the estimation of model parameters (including time delay) in a Smith predictor structure. *Technical Report EE/JVR/96/10*, Dublin City University, Dublin 9, Ireland.

O'Dwyer, A. (1996k). Model parameter and time delay estimation using a relay in place of the compensator in closed loop. *Technical Report EE/JVR/96/11*, Dublin City University, Dublin 9, Ireland.

O'Dwyer, A. (1996l). Implementation of other methods of model parameter and time delay estimation. *Technical Report EE/JVR/96/12*, Dublin City University, Dublin 9, Ireland.

O'Dwyer, A. (1996m). Development of theoretical results outlined in the thesis. *Technical Report EE/JVR/96/13*, Dublin City University, Dublin 9, Ireland.

Ogata, K. (1990). *Modern Control Engineering*, Prentice-Hall.

Ogunnaike, B.A. and Ray, W.H. (1979). Multivariable controller design for linear systems having multiple time delays. *AIChE Journal*, **25**, 1043-1057.

Ohta, T., Tanaka, H., Tanaka, K., Sannomiya, N. and Nishikawa, Y. (1979). A new optimisation method of PID control parameters for automatic tuning by process computer. *Proceedings of the IFAC Symposium: Computer Aided Design of Control Systems*, 133-138, Zurich, Switzerland.

Ortega, R. and Kelly, R. (1984). PID self-tuners: some theoretical and practical aspects. *IEEE Transactions on Industrial Electronics*, **IE-31**, 332-337.

Ortega, R. and Lozano, R. (1988). Globally stable adaptive controller for systems with delay. *International Journal of Control*, **47**, 17-23.

Owens, D.H. and Raya, A. (1982a). Robust stability of Smith predictor controllers for time-delay systems. *Research Report 179, Department of Control Engineering, University of Sheffield*, Mappin Street, Sheffield S1 3JD, England.

Owens, D.H. and Raya, A. (1982b). Robust stability of Smith predictor controllers for time-delay systems. *IEE Proceedings, Part D*, **129**, 298-304.

Ozbay, H. and Peery, T.E. (1993). On fixed order controllers for delay systems: discrete time case. *Proceedings of the American Control Conference*, 1030-1031, San Francisco, California, USA.

Ozturk, N. and Fardanesh, B. (1991). Multivariable control of the robot manipulator in the presence of time delay. *Proceedings of the International Conference on Industrial Electronics, Control and Instrumentation (IECON '91)*, Kobe, Japan, **2**, 1191-1193.

Pak, H.A. and Li, G.Q. (1992). Estimation of subsample delays for computer control systems. *Transactions of the ASME. Journal of Dynamic Systems, Measurement and Control*, **114**, 714-717.

Pal, J., Nagar, S.K. and Sharma, J.D. (1992). Digital controller design for systems with transport lag. *International Journal of Systems Science*, **23**, 2385-2392.

Palanisamy, K.R., Balachandran, K. and Ramasamy, R.S. (1988). Optimal control of linear time-varying delay systems via single-term Walsh series. *IEE Proceedings, Part D*, **135**, 332.

Pallas, M.-A. and Jourdain, G. (1991). Active high resolution time delay estimation for large BT signals. *IEEE Transactions on Signal Processing*, **39**, 781-787.

Palmor, Z.J. (1980). Stability properties of Smith dead-time compensator controllers. *International Journal of Control*, **32**, 937-949.

Palmor, Z.J. (1982). Properties of optimal stochastic control systems with dead-time. *Automatica*, **18**, 107-116.

Palmor, Z.J. and Halevi, Y. (1983). On the design and properties of multivariable dead time compensators. *Automatica*, **19**, 255-264.

Palmor, Z.J. and Powers, D.V. (1985). Improved dead-time compensator controllers. *AIChE Journal*, **31**, 215-221.

Palmor, Z.J. and Halevi, Y. (1990). Robustness properties of sampled-data systems with dead time compensators. *Automatica*, **26**, 637-640.

Palmor, Z.J. and Blau, M. (1994). An auto-tuner for Smith dead time compensator, *International Journal of Control*, **60**, 117-135.

Palmor, Z.J., Halevi, Y. and Krasney, N. (1995). Automatic tuning of decentralised PID controllers for TITO processes. *Automatica*, **31**, 1001-1010.

Pan, R. and Nikias, C.L. (1987). Phase reconstruction in the trispectrum domain. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-35**, 895-897.

Pandiscio, A.A. and Pearson, A.E. (1993). The reducing transformation and Smith predictor methods for stabilising plants with transport lag. *Proceedings of the American Control Conference*, 499-503, San Francisco, California, USA.

Papageorgiou, M. and Messner, A. (1989). Flow control of a long river stretch. *Automatica*, **25**, 177-183.

Paraskevopoulos, P.N. and Samiotakis, S.E. (1994). State space analysis and optimal control of time delay systems via Taylor's series. *IEE International Conference on Control (CONTROL '94),* 70-74.

Parrish, J.R. and Brosilow, C.B. (1985). Inferential control applications. *Automatica*, **21**, 527-538.

Patwardhan, A.A., Karim, M.N. and Shah, R. (1987). Controller tuning by a least-squares method. *AIChE Journal*, **33**, 1735-1737.

Pearson, A.E. and Wuu, C.Y. (1984). Decoupled delay estimation in the identification of differential delay systems. *Automatica*, **20**, 761-772.

Peebles, S.M., Hunter, S.R. and Corripio, A.N. (1994). Implementation of a dynamically-compensated PID control algorithm. *Computers and Chemical Engineering*, **18**, 995-1000.

Pemberton, T.J. (1972a). PID: The logical control algorithm. *Control Engineering*, **19, 5**, 66-67.

Pemberton, T.J. (1972b). PID: The logical control algorithm II. *Control Engineering*, **19, 7**, 61-63.

Penner, A. (1988). Tuning rules for a PI controller. *Proceedings of the ISA/88 International Conference and Exhibition*, 1037-1051, Houston, Texas, USA.

Perlmutter, D.D. (1965). *Introduction to Chemical Process Control*, John Wiley and Sons.

Perng, M.H. and Ju, J.S. (1994). Optimally decoupled robust control of MIMO plants with multiple delays. *IEE Proceedings - Control Theory and Applications*, **141**, 25-32.

Pessen, D.W. (1994). A new look at PID-controller tuning, *Transactions of the ASME. Journal of Dynamic Systems, Measurement and Control*, **116**, 553-557.

Peter, K. and Isermann, R. (1988). Predictive self-tuning control of a thermal process with unknown deadtimes. *Proceedings of the IFAC Adaptive Control of Chemical Processes Conference*, 69-74, Copenhagen, Denmark.

Peterka, V. (1989). Self-tuning control with alternative sets of uncertain process models. *Proceedings of the third IFAC Symposium on Adaptive Systems in Control and Signal Processing*, **2**, 409-414, Glasgow, U.K.

Petropulu, A., Nikias, C.L. and Proakis, J.G. (1988). Cumulant cepstrum of FM signals and high-resolution time delay estimation. *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, 2642-2645, New York, USA.

Philips, C.L. and Harbor, R.D. (1991). *Feedback control systems*, 2nd edition, Prentice-Hall.

Piche, R. (1990). Low-order rational all pass approximations to e^{-s}. *Proceedings of the 11th IFAC World Congress on Automatic Control*, **2**, 134-137, Tallinn, Estonia.

Pintelon, R. and Schoukens, J. (1990). Robust identification of transfer functions in the s and z domains. *IEEE Transactions on Instrumentation and Measurement*, **39**, 565-573.

Pintelon, R. and Van Biesen, L. (1990). Identification of transfer functions with time delay and its application to cable fault location. *IEEE Transactions on Instrumentation and Measurement*, **39**, 479-484.

Pintelon, R., Guillaume, P., Rolain, Y. and Verbeyst, F. (1992). Identification of linear systems captured in a feedback loop. *IEEE Transactions on Instrumentation and Measurement*, **41**, 747-754.

Pollard, A. (1971). *Process Control*, Heinemann Educational Books.

Poulin, E, Pomerleau, A. Desbiens, A. and Hodouin, D. (1996). Development and evaluation of an auto-tuning and adaptive PID controller. *Automatica*, **32**, 71-82.

Power, H.M. and Simpson, R.J. (1978). *Introduction to Dynamics and Control*, McGraw-Hill.

Prager, D.L. and Wellstead, P.E. (1980). Multivariable pole assignment self-tuning regulators. *Proceedings of the IEE, Part D*, **128**, 9-18.

Prasad, C.C., Hahn, V., Unbehauen, H. and Keuchel, U. (1985). Adaptive control of a variable dead time process with an integrator. *Proceedings of the IFAC Workshop on Adaptive Control of Chemical Processes*, 71-75, Frankfurt am Main, FRG.

Pratt, P. and Downing, C.J. (1994). LQG regulation of a thermal process - an alternative formulation. *Proceedings of the Irish DSP and Control Colloquium*, 177-184, Dublin City University, Dublin 9, Ireland.

Puleston, P.F. and Mantz, R.J. (1995). An anti-wind-up proportional integral structure for controlling time-delayed multiinput-multioutput processes. *Industrial and Engineering Chemistry Research*, **34**, 2993-3000.

Pupeikis, R. (1985). Recursive estimation of the parameters of linear systems with time delay. *Proceedings of the IFAC Identification and System Parameter Estimation Conference*, 787-792, York, U.K.

Qiu, G., Varley, M.R. and Terrell, T.J. (1992). Accelerated training of backpropagation networks by using adaptive momentum step. *Electronics Letters*, **28**, 377-379.

Rad, A.B. and Lo, W.L. (1994). Predictive PI controller. *International Journal of Control*, **60**, 953-975.

Rad, A.B., Tsang, K.M. and Lo, W.L. (1995). Adaptive control of dominant time delay systems via polynomial identification. *IEE Proceedings - Control Theory and Applications*, **142**, 433-438.

Radke, F. and Isermann, R. (1984). A parameter-adaptive PID-controller with stepwise parameter optimisation. *Proceedings of the IFAC 9th Triennial World Congress*, **4**, 1885-1890, Budapest, Hungary.

Rajakumar, A. and Krishnaswamy, P.R. (1975). Time to frequency domain conversion of step response data. *Industrial and Engineering Chemistry Process Design and Development*, **14**, 250-256.

Rake, H. (1980). Step response and frequency response methods. *Automatica*, **16**, 519-526.

Ralston, P.A.S., Watson, K.R., Patwardhan, A.A. and Deshpande, P.B. (1985). A computer algorithm for optimised control. *Industrial and Engineering Chemistry Process Design and Development*, **24**, 1132-1136.

Rangaiah, G.P. and Krishnaswamy, P.R. (1994a). Estimating second-order plus dead time model parameters. *Industrial and Engineering Chemistry Research*, **33**, 1967-1871.

Rangaiah, G.P. and Krishnaswamy, P.R. (1994b). Effective implementation of conservative model based control algorithm. *Chemical Engineering Communications*, **128**, 135-141.

Rangaiah, G.P. and Krishnaswamy, P.R. (1996). Estimating second-order dead time parameters from underdamped process transients. *Chemical Engineering Science*, **51**, 1149-1155.

Rangoussi, M. and Giannakis, G.B. (1991). FIR Modeling using log-bispectra: weighted least-squares algorithms and performance analysis. *IEEE Transactions on Circuits and Systems*, **38**, 281-296.

Rao, A.S. and Perdikaris, G.A. (1988). Computer aided compensator design. *IEEE Transactions on Education*, **31**, 221-224.

Rao, G.P. and Sivakumar, L. (1979). Identification of time-lag systems via Walsh functions. *IEEE Transactions on Automatic Control*, **AC-24**, 806-808.

Rao, G.P. and Palanisamy, K.R. (1983). Improved algorithms for parameter identification in continuous systems via Walsh functions. *IEE Proceedings, Part D*, **130**, 9-16.

Ray, W.H. and Soliman, M.A. (1970). The optimal control of processes containing pure time delays - I. Necessary conditions for an optimum. *Chemical Engineering Science*, **25**, 1911-1925.

Ray, W.H. (1981). *Advanced process control*, McGraw-Hill.

Readle, J.C. and Henry, R.M. (1994). On-line determination of time delay using multiple recursive estimators and fuzzy reasoning. *Proceedings of the International Conference on Control*, **2**, 1436-1441, Bradford, England.

Reddy, G.R. and Rao, V.V. (1987). Signal delay and waveform estimation through differential cepstrum averaging. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-35**, 1487-1489.

Reed, F.A., Feintuch, P.L. and Bershad, N.J. (1981). Time delay estimation using the LMS adaptive filter - static behaviour. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-29**, 561-571.

Reklaitis, G.V., Ravindran, A. and Ragsdell, K.M. (1983). *Engineering optimisation methods and applications*, John Wiley and Sons.

Ren, W. (1993). Stochastic adaptive control of multiple delay systems. *Proceedings of the American Control Conference*, 2195-2199, San Francisco, CA., USA.

Renders, J.-M. and Flasse, S.P. (1996). Hybrid methods using genetic algorithms for global optimisation. *IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernetics*, **26**, 243-258.

Ringwood, J.V. (1987). Algorithms for self tuning PID control - a survey. *Research Report: CTRU 8715*, School of Electronic Engineering, N.I.H.E., Dublin 9, Ireland.

Ringwood, J.V. and O'Dwyer, A. (1994a). A frequency domain based self tuning PID controller. *Proceedings of the Asian Control Conference*, 331-334, Tokyo, Japan.

Ringwood, J.V. and O'Dwyer, A. (1994b). A self-tuning PID controller based on on-line measurement of the process frequency response. *EURACO Workshop: 'Recent results in robust and adaptive control'*, Dublin University, Dublin 2, Ireland.

Rivera, D.E., Morari, M. and Skogestad, S. (1986). Internal Model Control. 4. PID controller design. *Industrial and Engineering Chemistry Process Design and Development*, **25**, 252-265.

Rivera, D.E. and Morari, M. (1990). Low-order SISO controller tuning methods for the H_2, H_∞ and μ objective functions. *Automatica*, **26**, 361-369.

Robinson, W.R. and Soudack, A.C. (1970). A method for the identification of time delays in linear systems. *IEEE Transactions on Automatic Control*, **AC-15**, 97-101.

Romagnoli, J.A., Karim, M.N., Agamennoni, O.E. and Desages, A. (1988). Controller designs for model-plant parameter mismatch. *IEE Proceedings, Part D*, **135**, 157-164.

Rovira, A.A., Murrill, P.W. and Smith, C.L. (1969). Tuning controllers for setpoint changes. *Instruments and Control Systems*, **December**, 67-69.

Rovira, A.A., Murrill, P.W. and Smith, C.L. (1970). Modified PI algorithm for digital control. *Instruments and Control Systems*, **August**, 101-102.

Roy, S., Malik, O.P. and Hope, G.S. (1990). Predictors for application to real-time adaptive control of a diesel prime-mover. *Proceedings of the Annual meeting of the IEEE Industry Applications Society*, **2**, 1804-1810.

Roy, S., Malik, O.P. and Hope, G.S. (1991a). A least-squares based model-fitting identification technique for diesel prime-movers with unknown dead-time, *IEEE Transactions on Energy Conversion*, **6**, 251-256.

Roy, S., Malik, O.P. and Hope, G.S. (1991b). Adaptive control of plants using all-zero model for dead-time identification. *IEE Proceedings, Part D*, **138**, 445-452.

Roy, S., Malik, O.P. and Hope, G.S. (1991c). An adaptive control scheme for speed control of diesel driven power plants. *IEEE Transactions on Energy Conversion*, **6**, 605-611.

Roy, S., Malik, O.P. and Hope, G.S. (1993a). A k-step predictive scheme for speed control of diesel driven power plants. *IEEE Transactions on Industry Applications*, **29**, 389-396.

Roy, S., Malik, O.P. and Hope, G.S. (1993b). Real-time test results with adaptive speed controllers for a diesel prime-mover. *IEEE Transactions on Energy Conversion*, **8**, 499-505.

Ruano, A.E.B., Fleming, P.J. and Jones, D.I. (1992). Connectionist approach to PID autotuning. *IEE Proceedings, Part D*, **139**, 279-285.

Salgado, M.E., De Souza, C.E. and Goodwin, G.C. (1988). Issues in time delay modelling. *Proceedings of the Eighth IFAC/IFORS Symposium on Identification and System Parameter Estimation*, 549-554, Beijing, PRC.

Salt, J.E., Daku, B.L.F. and Wood, H.C. (1993). Noncontact velocity measurement using time delay estimation techniques. *IEEE Transactions on Signal Processing*, **41**, 288-295.

Sanathanan, C.K. and Quinn, S.B. (1987). Design of set point regulators for processes involving time delay. *AIChE Journal*, **33**, 1873-1881.

Sandoz, D.J. (1987). Practical experience of the control of industrial systems with time delay. *Report, Vumen Ltd.*, Industrial Control Systems Division, Enterprise House, Lloyd Street North, Manchester, MI5 E4N, England.

Santacesaria, C. and Scattolini, R. (1993). Easy tuning of Smith predictor in presence of delay uncertainty. *Automatica*, **29**, 1595-1597.

Sasaki, K., Sato, T. and Makamura, Y. (1977). Holographic passive sonar. *IEEE Transactions on Sonics and Ultrasonics*, **SU-24**, 193-200.

Sato, A. (1991). An analytical study of the momentum term in a backpropagation algorithm. *Proceedings of the 1991 International Conference on Artificial Neural Networks (ICANN-'91),* **1**, 617-622.

Sato, T. and Sasaki, K. (1977). Bispectral holography. *Journal of the Acoustical Society of America*, **62**, 404-408.

Scales, L.E. (1985). *Introduction to non-linear optimisation*, Macmillan Publishers Ltd.

Scarborough, K., Ahmed, N. and Carter, G.C. (1981). On the simulation of a class of time delay estimation algorithms. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-29**, 534-540.

Scattolini, R. (1986). A multivariable self-tuning controller with integral action. *Automatica*, **22**, 619-627.

Schei, T.S. (1992). A method for closed loop automatic tuning of PID controllers. *Automatica*, **28**, 587-591.

Schei, T.S. (1994). Automatic tuning of PID controllers based on transfer function estimation. *Automatica*, **30**, 1983-1989.

Schneider, D.M. (1988). Control of processes with time delays. *IEEE Transactions on Industry Applications*, **24**, 186-191.

Schoukens, J., Pintleton, R. and Renneboog, J. (1988). A maximum likelihood estimator for linear and non linear systems - a practical application of estimation techniques in measurement problems. *IEEE Transactions on Instrumentation and Measurement*, **37**, 10-17.

Schuster, T. (1989). An adaptive PID-controller. *Periodica Polytechnica, Electrical Engineering*, **33**, 263-272.

Schwartzenbach, J. and Gill, K.F. (1984). *System Modelling and Control*, Second Edition, Edward Arnold.

Seborg, D.E., Edgar, T.F. and Shah, S.L. (1986). Adaptive control strategies for process control: a survey. *AIChE Journal*, **32**, 881-913.

Seborg, D.E., Edgar, T.F. and Mellichamp, D.A. (1989). *Process dynamics and control*, John Wiley and Sons.

Segal, M., Weinstein, E. and Musicus, B.R. (1991). Estimate-maximize algorithms for multichannel time delay and signal estimation. *IEEE Transactions on Signal Processing*, **39**, 1-16.

Semino, D. and Scali, C. (1994). A method for robust tuning of linear quadratic optimal controllers. *Industrial and Engineering Chemistry Research*, **33**, 889-895.

Shafiei, Z. and Shenton, A.F. (1994). Tuning of PID-type controllers for stable and unstable systems with time delay. *Automatica*, **30**, 1609-1615.

Shah, S.A., Woodruff, E.A. and Northrop, R.B. (1988). An adaptive IPFM/SDC controller for the closed-loop sodium nitroprusside regulation of post-operative blood pressure. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, **2**, 519-520, New Orleans, LA, USA.

Shahian, B. and Hassul, M. (1993). *Control system design using MATLAB*, Prentice-Hall.

Shahrohki, M. and Naimpour, F. (1992). A feedback control scheme for SISO systems with time delays and constrained inputs. *Computers and Chemical Engineering*, **16**, 1073-1076.

Shaw, J.A. (1993). Analysis of traditional PID tuning methods. *Proceedings of the ISA/93 Advances in Instrumentation and Control Conference*, **48**, 391-397, McCormick Place, Chicago, Illinois, USA.

Shen, S.-H. and Yu, C.-C. (1994). Use of relay-feedback test for automatic tuning of multivariable systems. *AIChE Journal*, **40**, 627-646.

Shen, S.-H., Yu, H.-D. and Yu, C.-C. (1996a). Use of saturation-relay feedback for autotune identification. *Chemical Engineering Science*, **51**, 1187-1198.

Shen, S.-H., Wu, J.-S. and Yu, C.-C. (1996b). Use of biased-relay feedback for system identification. *AIChE Journal*, **42**, 1174-1180.

Shen, S.-H., Wu, J.-S. and Yu, C.-C. (1996c). Autotune identification under load disturbance. *Industrial and Engineering Chemistry Research*, **35**, 1642-1651.

Shen, Y.P., Rangayyan, R.M., Frank, C.B., Bell, G.D. and Zhang, Y.D. (1993). Time delay estimation for source localisation of vibroarthrographic signals from human knee joints. *Proceedings of the Annual Conference on Engineering in Medicine and Biology*, **15**, 381-382.

Sheridan, T.B. (1993). Space teleoperation through time delay: review and prognosis. *IEEE Transactions on Robotics and Automation*, **9**, 592-606.

Shigemasa, T. and Akizuki, K. (1981). A closed loop auto-tuning method for digital PID controller. *Proceedings of the IEEE/IECI Conference on applications of mini and microcomputers*, 427-432, San Francisco, USA.

Shinskey, F.G. (1990). How good are our controllers in absolute performance and robustness?. *Measurement and Control*, **23**, 114-121.

Shu, Z., Watanabe, K. and Yamade, K. (1994). Robust internal model control of time-delay system. *Proceedings of the Asian Control Conference*, 1045-1048, Tokyo, Japan.

Silva, F.M. and Almeida, L.B. (1990). Speeding up backpropagation. In R. Eckmiller (Ed.), *Advanced Neural Computers*, 151-158, Elsevier Science Publishers B.V. (North Holland).

Silva, M.T. (1987). Time delay estimation. In D.F. Elliott (Ed.), *Handbook of Digital Signal Processing Engineering Applications*, Academic Press Inc.

Singh, A. and McEwan, D.H. (1976). The control of a process having appreciable transport lag - a laboratory case study. *IEEE Transactions on Industrial Electronics and Control Instrumentation*, **IECI-22**, 396-401.

Smirthwaite, P.T., Fisher, A.C., Gaylor, J.D., Henderson, I.A. and McGhee, J. (1994). Frequency domain identification of a blood urea nitrogen sensor with multifrequency binary signals. *IEE International Control Conference (CONTROL '94)*, 314-319.

Smith, C.A. and Corripio, A.B. (1985). *Principles and practice of automatic process control*, John Wiley and Sons.

Smith, J.O. and Friedlander, B. (1985). Adaptive interpolated time-delay estimation. *IEEE Transactions on Aerospace and Electronic Systems*, **AES-21**, 180-199.

Smith, O.J.M. (1957). Closer control of loops with dead time. *Chemical Engineering Progress*, **53**, 217-219.

So, H.C. and Ching, P.C. (1993). Target localisation in the presence of multipaths. *Electronics Letters*, **29**, 293-294.

So, H.C., Ching, P.C. and Chan, Y.T. (1994). A new algorithm for explicit adaptation of time delay. *IEEE Transactions on Signal Processing*, **42**, 1816-1820.

Soderstrom, T. and Stoica, P. (1989). *System Identification*, Prentice-Hall.

Soeterboek, R. (1992). *Predictive Control - a unified approach*, Prentice-Hall Inc.

Song, J.-B. and Hardt, D.E. (1994). Dynamic modeling and adaptive control of the gas-metal arc welding process. *Transactions of the ASME. Journal of Dynamic Systems, Measurement and Control*, **116**, 405-413.

Song, W.-Z. and Xu, Y. (1985). Structure and parameter identification for a kind of multivariable linear systems with unknown time delays. *Proceedings of the IFAC Identification and System Parameter Estimation Conference*, 793-798, York, U.K.

Stahl, H. and Hippe, P. (1987). Design of pole placing controllers for stable and unstable systems with pure time delay. *International Journal of Control*, **45**, 2173-2182.

Stein, S. (1981). Algorithms for ambiguity function processing. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-29, II**, 588-599.

Stoica, P. (1983). Generalised Yule-Walker equations and testing the orders of multivariate time series. *International Journal of Control*, **37**, 1159-1166.

Stoica, P., Eykhoff, P., Janssen, P. and Soderstrom, T. (1986). Model-structure selection by cross validation. *International Journal of Control*, **43**, 1841-1878.

Su, L.-S. (1993). Applications of digital controller in process control. *Proceedings of the ISA/93 Advances in Instrumentation and Control Conference*, **48**, 625-632, McCormick Place, Chicago, Illinois, USA.

Sundaresan, K.R., Chandra Prasad, C. and Krishnaswamy, P.R. (1978). Evaluating parameters from process transients. *Industrial and Engineering Chemistry Process Design and Development*, **17**, 237-241.

Sundaresan, K.R. and Krishnaswamy, P.R. (1978). Estimation of time delay, time constant parameters in time, frequency and Laplace domains. *The Canadian Journal of Chemical Engineering*, **56**, 257-262.

Sung, S.J., Park, H.I., Lee, I.-L. and Yang, D.R. (1994). On-line process identification and autotuning using P-controller. *Proceedings of the Asian Control Conference*, 411-414, Tokyo, Japan.

Tachibana, Y. (1984). Identification of a system with dead time and its application to auto tuner. *Electrical Engineering in Japan*, **104**, 128-137.

Tade, M.O., Bayoumi, M.M. and Bacon, D.W. (1988a). Self-tuning controller design for systems with arbitrary time delays. Part 1: Theoretical development. *International Journal of System Science*, **19**, 1095-1115.

Tade, M.O., Bayoumi, M.M. and Bacon, D.W. (1988b). Self-tuning controller design for systems with arbitrary time delays. Part 2: Algorithms and simulation examples. *International Journal of System Science*, **19**, 1117-1141.

Tan, Y. and De Keyser, R. (1994). Auto-tuning PID control using neural predictor to compensate large time-delay. *Proceedings of the Third IEEE Conference on Control Applications*, **2**, 1429-1434, Glasgow, U.K.

Tan, Y.-L. and Weber, T.W. (1985). Controller tuning of a third order process under proportional-integral control. *Industrial and Engineering Chemistry Process Design and Development*, **24**, 1155-1160.

Teng, F.-C. and Sirisena, H.R. (1988). Self-tuning PID controllers for dead time processes. *IEEE Transactions on Industrial Electronics*, **35**, 119-125.

Teng, F.-C. (1990). Investigation of the enlarged least-square estimator combined with a classical PID controller. *Transactions of the Institute of Measurement and Control*, **12**, 224-228.

Teng, F.C., Ledwich, G.F. and Tsoi, A.C. (1994). Extension of the Dahlin-Higham controller to multivariable systems with time delays. *International Journal of System Science*, **25**, 337-350.

Terrett, P. and Downing, C.J. (1993). System identification and modelling on the TMS320C25 fixed-point processor using the delta operator. *Irish Colloquium on DSP and Control*, 239-247, University College Dublin, Dublin 4, Ireland.

Terrett, P. and Downing, C.J. (1994). Parameter estimation on a fixed point processor using the delta operator. *Transactions of the Institute of Measurement and Control*, **16**, 163-172.

Thomson, M., Cassidy, P.G. and Sandoz, D.J. (1989). Automatic tuning of PID controllers using a combined time- and frequency-domain method. *Transactions of the Institute of Measurement and Control*, **11**, 40-47.

Thompson, S. (1993). Internal model control and feedback systems. *Irish Colloquium on DSP and Control*, 133-140, University College Dublin, Dublin 4, Ireland.

Tjokro, S. and Shah, S.L. (1985). Adaptive PID control. *Proceedings of the American Control Conference*, **3**, 1528-1534.

Tremblay, R.J., Carter, G.C. and Lytle, D.W. (1987). A practical approach to the estimation of amplitude and time-delay parameters of a composite signal. *The Journal of Oceanic Engineering*, **OE-12**, 273-278.

Triantafyllou, M.S. and Grosenbaugh, M.A. (1991). Robust control for underwater vehicle systems with time delays, *IEEE Journal on Oceanic Engineering*, **16**, 146-151.

Tsai, J.S.H., Wei, C.-P. and Shieh, L.S. (1994). Design of a stable digital output feedback controller for continuous-time input-delay plants. *International Journal of Systems Science*, **25**, 2187-2206.

Tse, E. and Weinert, H.L. (1975). Structure determination and parameter identification for multivariable stochastic linear systems. *IEEE Transactions on Automatic Control*, **AC-20**, 603-613.

Tuch, J., Feuer, A. and Palmor, Z.J. (1994). Time delay estimation in continuous linear time invariant systems. *IEEE Transactions on Automatic Control*, **39**, 823-827.

Tugnait, J.K. (1991). On time delay estimation with unknown spatially correlated Gaussian noise using fourth-order cumulants and cross cumulants. *IEEE Transactions on Signal Processing*, **39**, 1258-1267.

Tyreus, B.D. and Luyben, W.L. (1992). Tuning PI controllers for integrator/dead time processes. *Industrial and Engineering Chemistry Research*, **31**, 2625-2628.

Unbehauen, H. and Gohring, B. (1974). Tests for determining model order in parameter estimation. *Automatica*, **10**, 233-244.

Unbehauen, H. and Rao, G.P. (1987). *Identification of continuous systems*, North Holland System and Control Series, Volume 10, Elsevier Science Publishers B.V.

Uraz, A. and Ozturk, N. (1985). Predictor control of multivariable processes with time delays. *IEE Control Conference 1985*, 648-653.

Van den Boom, A.J.W. and Van den Enden, A.W.M. (1974). The determination of the orders of process and noise dynamics. *Automatica*, **10**, 245-256.

Vanderplaats, G.N. (1984). *Numerical optimisation techniques for engineering design: with applications*, McGraw-Hill.

Vandeursen, J.M. and Peperstraete, J.A. (1995). Internal model control with improved disturbance rejection. *International Journal of Control*, **62**, 983-999.

Van Overbeek, A.J.M. and Ljung, L. (1982). On-line structure selection for multivariable state-space models. *Automatica*, **18**, 529-543.

Vasilev, V.N. and Aidemirski, P.P. (1990). Time delay estimation with gradient type adaptive algorithm. *Electronics Letters*, **26**, 1797-1798.

Venkatashankar, V. and Chidambaram, M. (1994). Design of P and PI controllers for unstable first order plus time delay systems. *International Journal of Control*, **60**, 137-144.

Vermeer, P.J., Morris, A.J. and Shah, S.L. (1988). Adaptive PID control - a pole placement algorithm with a single controller tuning parameter. *Proceedings of the IFAC Adaptive Control of Chemical Processes Conference*, 159-164, Copenhagen, Denmark.

Vit, K. (1979). Smith-like predictor for control of parameter distributed processes. *International Journal of Control*, **31**, 179-193.

Voda, A., Landau, I.D., M'Saad, M. and Bouslimani, M. (1994). Industrial evaluation of the KLV method for the auto-calibration of PID controllers. *IEE International Control Conference (CONTROL '94)*, 1507-1511.

Voda, A. and Landau, I.D. (1995a). A method for the auto-calibration of PID controllers. *Automatica*, **30**, 41-53.

Voda, A. and Landau, I.D. (1995b). The autocalibration of PI controllers based on two frequency measurements. *International Journal of Adaptive Control and Signal Processing*, **9**, 395-421.

Vogel, E.F. and Edgar, T.F. (1982). Application of an adaptive pole-zero placement controller to chemical processes with variable dead time. *Proceedings of the American Control Conference*, 536-544.

Vogel, E.F. and Edgar, T.F. (1988). An adaptive pole placement controller for chemical processes with variable dead time. *Computers and Chemical Engineering*, **12**, 15-26.

Voss, G.I., Chizeck, H.J. and Katona, P.G. (1988). Self-tuning controller for drug delivery systems. *International Journal of Control*, **47**, 1507-1520.

Wakeland, W.R. (1976). Bode compensator design. *IEEE Transactions on Automatic Control*, **October**, 771-773.

Wang, F.-S. (1990). Adaptive root locus control for SISO processes with time delays. *Optimal Control Applications and Methods*, **11**, 211-221.

Wang, F.-S., Juang, W.-S. and Chan, C.-T. (1993a). Optimal tuning of cascade PID control systems. *Proceedings of the Second IEEE Conference on Control Applications* **2**, 825-828, Vancouver, B.C., Canada.

Wang, F.-S., Juang, W.-S. and Chan, C.-T. (1995a). Optimal tuning of PID controllers for single and cascade control loops. *Chemical Engineering Communications*, **132**, 15-34.

Wang, F.-S. and Wu, T.-Y. (1995). Multiple loop PID controller tuning by the goal attainment trade-off method. *Transactions of the Institute of Measurement and Control*, **17**, 27-34.

Wang, J. and Wan, B.-W. (1988). Control approaches for processes with large time delays: a review. *ISA Transactions*, **27**, 61-65.

Wang, Q.-G., Sun, Y.X. and Zhou, C.H. (1988). Finite spectrum assignment for multivariable delay systems in the frequency domain. *International Journal of Control*, **47**, 729-734.

Wang, Q.-G., Lee, T.H. and Tan, K.K. (1995b). Automatic tuning of finite spectrum assignment controllers for delay systems. *Automatica*, **31**, 477-482.

Wang, S.-S. and Chen, B.-S. (1988). Optimal model-matching control for time-delay systems. *International Journal of Control*, **47**, 883-894.

Wang, T.-S. and Clements, W.C. (1995). Adaptive multivariable PID control of a distillation column with unknown and varying dead time. *Chemical Engineering Communications*, **132**, 1-13.

Wang, Z.-Q. and Skogestad, S. (1993). Robust control of time delay systems using the Smith predictor. *International Journal of Control*, **57**, 1405-1420.

Wang, Z.-Q., Lundstrom, P. and Skogestad, S. (1994). Representation of uncertain time delays in the H infinity framework. *International Journal of Control*, **59**, 627-638.

Warwick, K. and Kang, Y.H. (1993). Genetic algorithms for on-line system identification. *Proceedings of the International Conference on Artificial Neural Networks and Genetic Algorithms*, 436-440, Springer-Verlag.

Watanabe, K. and Ito, M. (1981). A process model control for linear systems with delays. *IEEE Transactions on Automatic Control*, **AC-26**, 1261-1269.

Watanabe, K., Ishiyama, Y. and Ito, M. (1983). Modified Smith predictor control for multivariable systems with delays and unmeasurable step disturbances. *International Journal of Control*, **73**, 959-973.

Watanabe, K. and Sato, M. (1984). A process model control for multivariable systems with multiple delays in inputs and outputs subject to unmeasurable disturbances. *International Journal of Control*, **39**, 1-17.

Watanabe, K. and Ouchi, T. (1985). An observer of systems with delays in state variables. *International Journal of Control*, **41**, 217-229.

Weber, T.W. and Bhalodia, M. (1979). Optimum behaviour of a third order process under feedback control. *Industrial and Engineering Chemistry Process Design and Development*, **18**, 217-223.

Webster, D. (1994). A pulsed ultrasonic distance measurement system based on phase digitising. *IEEE Transactions on Instrumentation and Measurement*, **43**, 578-582.

Weerasooriya, S. and Phan, D.T. (1995). Discrete time LQG/LTR design and modeling of a disk drive actuator tracking servo system. *IEEE Transactions on Industrial Electronics*, **42**, 240-247.

Weiss, A.J. and Stein, Z. (1987). Optimal below threshold delay estimation for radio signals. *IEEE Transactions on Aerospace and Electronic Systems*, **AES-23**, 726-730.

Wellons, M.C. and Edgar, T.F. (1987). The generalised analytical predictor. *Industrial and Engineering Chemistry Research*, **26**, 1523-1536.

Wellstead, P.E. (1978). An instrumental product moment test for model order estimation. *Automatica*, **14**, 89-91.

Wellstead, P.E., Prager, D. and Zanker, P. (1979). Pole assignment self-tuning regulator. *IEE Proceedings*, **126**, 781-787.

Wellstead, P.E. (1981). Non parametric methods of system identification. *Automatica*, **17**, 55-69.

Wellstead, P.E. (1986). *Lecture notes in control and information sciences, Vol. 79: Signal Processing for Control.* Editors: K. Godfrey, P. Jones. Series Editors: M. Thoma, A. Wyner. Springer-Verlag.

Wellstead, P.E. and Zarrop, M.B. (1991). *Self-tuning systems.* John Wiley and Sons.

Whalley, R. and Zeng, Z. (1994). Mismatch Smith predictors. *Transactions of the Institute of Measurement and Control*, **16**, 174-182.

Whalley, R. and Zeng, Z. (1996). A robust Smith predictor. *Transactions of the Institute of Chemical Engineers*, **74**, **A**, 21-29.

Whitfield, A.H. (1986). Transfer function synthesis using frequency response data. *International Journal of Control*, **43**, 1413-1426.

Whitfield, A.H. (1987). Asymptotic behaviour of transfer function synthesis methods. *International Journal of Control*, **45**, 1083-1092.

Widrow, B. and Stearns, S.D. (1985). *Adaptive Signal Processing*, Prentice-Hall Inc.

Wilkinson, A.J., Roberts, S., Taylor, P.M. and Taylor, G.E. (1991). The application of recursive identification based on delta operators. *International Conference on Control,* **1**, 627-631, Edinburgh, Scotland.

Willis, M.J., Montague, G.A., Di Massimo, C., Tham, M.T. and Morris, A.J. (1992). Artificial neural networks in process estimation and control. *Automatica*, **28**, 1181-1187.

Wittenmark, B. and Astrom, K.J. (1984). Practical issues in the implementation of self tuning controllers. *Automatica*, **20**, 595-605.

Wong, E.C. (1980). Stochastic systems with time delay. In C.T. Leondes, *Control and Dynamic Systems. Advances in Theory and Applications*, Vol. 16, pp. 131-176, Academic Press.

Wong, K.Y. and Bayoumi, M.M. (1982). A self-tuning control algorithm for systems with unknown time delay. *Proceedings of the IFAC Identification and System Parameter Estimation Conference*, 1193-1198, Washington, D.C., U.S.A.

Wong, S.K.P. and Seborg, D.E. (1986). A theoretical analysis of Smith and analytical predictors. *AIChE Journal*, **32**, 1597-1603.

Wu, H., Mizukami, K. and Willgoss, R.A. (1993). Decentralized robust control for uncertain large scale time-delay dynamical systems. *Control - Theory and Advanced Technology*, **9**, 533-546.

Wu, W.-T. and Tseng, C.-G. (1992). Robust and adaptive control of a multi-variable system. *Transactions of the Institute of Measurement and Control*, **14**, 204-211.

Wu, W.-T., Tseng, C.-G. and Chu, Y.T. (1994). System identification and on-line robust control of a multivariable system. *International Journal of System Science*, **25**, 423-439.

Wuu, C.-Y. and Pearson, A.E. (1984). On time delay estimation involving received signals. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-32**, 828-835.

Xia, L., Moore, J.B. and Gevers, M. (1987). On adaptive estimation and pole assignment of overparametrized systems. *International Journal of Adaptive Control and Signal Processing*, **1**, 143-160.

Xia, L. and Moore, J.B. (1989). Recursive identification of overparameterised systems. *IEEE Transactions on Automatic Control*, **34**, 327-331.

Xu, C.-W. (1989). Fuzzy systems identification. *IEE Proceedings, Part D*, **136**, 146-150.

Xu, H. (1988). Robust protection-time delay tracing adaptive control and its application. *Proceedings of the Eighth IFAC/IFORS Symposium on Identification and System Parameter Estimation*, 769-774, Beijing, PRC.

Yamada, H., Ohmiya, M., Ogawa, Y. and Itoh, K. (1991). Superresolution techniques for time-domain measurements with a network analyzer. *IEEE Transactions on Antennas and Propagation*, **39**, 177-183.

Yamamoto, T., Ishihara, H., Omatu, S. and Kitamori, T. (1991). Multivariable self tuning controller with I-PD structure. *Proceedings of the 1991 International Conference on Industrial Electronics, Control and Instrumentation (IECON '91)*, **3**, 1812-1817, Kobe, Japan.

Yamamoto, T., Omatu, S. and Kaneda, M. (1994). A design method for self-tuning PID controllers. *Proceedings of the American Control Conference*, **3**, 3263-3267.

Yamanaka, K. and Shimemura, E. (1987). Effects of mismatched Smith controller on stability in systems with time-delay. *Automatica*, **23**, 787-791.

Yang, J., Chen, C.S. and Xu, Y. (1994). Automatic tuning of phase-lead and phase-lag compensators. *International Journal of Control*, **60**, 631-640.

Yang, J.-S. (1993). A parameter plane method for the PID control of a multirate, sampled-data chemical reactor system with a transportation lag. *Chemical Engineering Communications*, **122**, 227-244.

Yang, T.C. (1994). On computational delay in digital and adaptive controllers. *IEE International Control Conference (CONTROL '94)*, 906-909.

Yang, Z.-J., Hachino, T. and Tsuji, T. (1996). Model reduction with time delay combining the least-squares method with the genetic algorithm. *IEE Proceedings - Control Theory and Applications*, **143**, 247-254.

Yao, Y.X., Zhang, Y.M. and Kovacevic, R. (1996). Parameterisation of observers for time delay systems and its application to observer design. *IEE Proceedings - Control Theory and Applications*, **143**, 225-231.

Yastrebov, A. and Grzywaczewski, M. (1994). Identification of linear continuous systems with time delays on the basis of discrete measurement data. *Proceedings of the Asian Control Conference*, 929-932, Tokyo, Japan.

Ydstie, B.E. (1984). Extended horizon adaptive control. *Proceedings of the 9th IFAC World Congress*, 911-915, Budapest, Hungary.

Yin, K. and Asbjornson, O.A. (1993). Application of MIMO adaptive control to a binary distillation process. *Chemical Engineering Communications*, **124**, 115-130.

Youn, D.H., Ahmed, N. and Carter, G.C. (1982). On using the LMS algorithm for time delay estimation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-30**, 798-801.

Youn, D.H., Ahmed, N. and Carter, G.C. (1983). On the Roth and SCOT algorithms: time domain implementations. *Proceedings of the IEEE*, **71**, 536-538.

Youn, D.H. and Mathews, V.J. (1984). Adaptive realizations of the maximum likelihood processor for time delay estimation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-32**, 938-940.

Young, C., O'Malley, M. and De Paor, A.M. (1990). Control of reverse-reaction processes via a pseudo-predictor. *International Journal of Control*, **42**, 261-264.

Young, G.E., Rao, K.S.S and Chatufale, V.R. (1995). Block-recursive identification of parameters and delay in the presence of noise. *Transactions of the ASME. Journal of Dynamic Systems, Measurement and Control*, **117**, 600-607.

Yusof, R., Omatu, S. and Khalid, M. (1994). Temperature control of a water bath by self-tuning PI controller. *International Journal of Systems Science*, **25**, 1391-1404.

Yuwana, M. and Seborg, D.E. (1982). A new method for on line controller tuning. *AIChE Journal*, **28**, 434-440.

Zafiriou, E. and Morari, M. (1985). Digital controllers for SISO systems: a review and a new algorithm. *International Journal of Control*, **42**, 885-876.

Zevros, C.C. and Dumont, G.A. (1988). Deterministic adaptive control based on Laguerre series representation. *International Journal of Control*, **48**, 2333-2359.

Zevros, C., Belanger, P.R. and Dumont, G.A. (1988). On PID controller tuning using orthonormal series identification. *Automatica*, **24**, 165-175.

Zhang, J.-F. and Chen, H.-F. (1990). Identification of coefficients, orders and time-delay for ARMAX systems. *Proceedings of the 11th IFAC World Congress on Automatic Control*, **3**, 129-134, Tallinn, Estonia.

Zhang, W. and Raghuveer, M. (1991). Nonparametric bispectrum-based time-delay estimators for multiple sensor data. *IEEE Transactions on Signal Processing*, **39**, 770-774.

Zhang, Z.-J., Li, D.-F. and Yuan, T.-X. (1985). On line identification of MIMO linear discrete stochastic systems. *Proceedings of the IFAC Identification and System Parameter Estimation Conference*, 1495-1490, York, U.K.

Zhao, Z.-Y. and Sagara, S. (1990). Consistent estimation of time delay in continuous systems. *Proceedings of the 29th SICE Annual Conference (SICE '90)*, **11**, 753-756.

Zhao, Z.-Y., Tomikuza, M. and Iskra, S. (1993). Fuzzy gain scheduling of PID controllers. *IEEE Transactions on Systems, Man and Cybernetics*, **23**, 1392-1398.

Zheng, F., Cheng, M. and Cao, W.-B. (1995). Variable structure control of time-delay systems with a simulation study on stabilising combustion in liquid propellant rocket motors. *Automatica*, **31**, 1031-1037.

Zheng, W.-X. and Feng, C.-B. (1988). Unbiased identification of linear time-delay systems corrupted by correlated noise. *Proceedings of the IFAC/IFORS Symposium on Identification and System Parameter Estimation*, 477-482, Beijing, PRC.

Zheng, W.-X. and Feng, C.-B. (1990). Hybrid method for model reduction with time delay. *International Journal of Systems Science*, **21**, 755-763.

Zhu, J.J-M. and Saucier, M. (1992). An IMC-based extended PID controller for sampled data systems, *Proceedings of the American Control Conference*, **1**, 601-606, Chicago, IL., USA.

Zhu, K.Y. and Ling, K.V. (1994). Model-based predictive controller. *IEE International Control Conference (CONTROL '94)*, 1012-1016.

Zhuang, M. and Atherton, D.P. (1993). Automatic tuning of optimum PID controllers. *IEE Proceedings, Part D*, **140**, 216-224.

Zhuang, M. and Atherton, D.P. (1994a). PID controller design for a TITO system. *IEE Proceedings - Control Theory and Applications*, **141**, 111-120.

Zhuang, M. and Atherton, D.P. (1994b). Optimum cascade PID controller design for SISO systems. *IEE International Control Conference (CONTROL '94)*, 606-611.

Ziegler, J.G. and Nichols, N.B. (1942). Optimum settings for automatic controllers. *Transactions of the ASME*, **November**, 759-768.

Ziegler, J.G. and Nichols, N.B. (1943). Process lags in automatic control circuits. *Transactions of the ASME*, **July**, 433-444.